

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2004, 01:28:51 ; Search time 154 Seconds

(without alignments)
444.918 Million cell updates/sec

Title: US-09-084-691b-206

Perfect score: 1045

Sequence: 1 MSTLEPKPORKTKENTNRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 23Sep04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1045	100.0	191	2	AAR92987 Hepatitis
2	996	95.3	191	2	AAR92972 Hepatitis
3	982	94.0	191	2	AAR92974 Hepatitis
4	982	94.0	191	2	AAR92973 Hepatitis
5	981	93.9	191	2	AAR92953 Hepatitis
6	980	93.8	191	2	AAR92978 Hepatitis
7	978	93.6	191	2	AAR92976 Hepatitis
8	977	93.5	191	2	AAR92977 Hepatitis
9	977	93.5	196	2	AAR74048 Synthetic
10	977	93.5	196	2	AAR74047 Synthetic
11	977	93.5	319	2	AAR96546 Hepatitis
12	977	93.5	326	2	AAR922137 HCV-HC59
13	977	93.5	733	2	AAR38278 HCV
14	977	93.5	2894	2	AAR24440 Composite
15	977	93.5	2894	2	AAR70230 Composite
16	977	93.5	3011	2	AAR66995 Hepatitis
17	976	93.4	191	2	AAR44010 Hepatitis
18	976	93.4	191	2	AAR92938 Hepatitis
19	976	93.4	191	2	AAR92941 Hepatitis
20	976	93.4	191	2	AAR92937 Hepatitis
21	976	93.4	191	2	AAR92939 Hepatitis
22	976	93.4	191	2	AAR92940 Hepatitis
23	976	93.4	191	3	AAY94411 Human hep
24	976	93.4	249	8	ADJ10436 HCV Core
25	976	93.4	473	8	ADJ10438 HCV Core (

ALIGNMENTS

RESULT 1

AAR92987

ID AAR92987 standard; protein; 191 AA.

XX AC AAR92987;

XX AC AAR92987;

DT 02-OCT-1996 (first entry)

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

XX DE Hepatitis C virus isolate HK2 core protein.

26	976	93.4	502	2	AAR67591	Hepatitis
27	976	93.4	967	2	AAR79222	pHCV141-e
28	976	93.4	1006	2	AAW12715	HCV genom
29	976	93.4	1648	2	AAR79221	pHCV176-e
30	976	93.4	2984	4	AAE00449	Hepatitis
31	976	93.4	2984	4	AAE00447	Hepatitis
32	976	93.4	2984	4	AAE00442	Hepatitis
33	976	93.4	3002	7	ADM24822	Hepatitis
34	976	93.4	3011	2	AAR40119	HCV genom
35	976	93.4	3011	2	AAR40120	HCV genom
36	976	93.4	3011	2	AAR79232	HCV sequ
37	976	93.4	3011	2	AAW77397	Hepatitis
38	976	93.4	3011	2	AAW77398	Hepatitis
39	976	93.4	3011	2	AAW98020	Infectiou
40	976	93.4	3011	4	AAE59173	Protein e
41	976	93.4	3011	4	AAE31169	Amino aci
42	976	93.4	3011	5	AAU84597	HCV polyp
43	976	93.4	3011	5	AAU79221	Hepatitis
44	976	93.4	3011	5	AAE19888	Hepatitis
45	976	93.4	3011	6	ABP71460	Amino aci

Query Match

Best Local Similarity

Score 1045; DB 2; Length 191;

Pred. No. 2.5e-90;

Sequence 191 AA;

Query Match

Best Local Similarity

Score 1045; DB 2; Length 191;

Pred. No. 2.5e-90;

Sequence 191 AA;

Query Match

Best Local Similarity

Score 1045; DB 2; Length 191;

Pred. No. 2.5e-90;

Sequence 191 AA;

Query Match

Best Local Similarity

Score 1045; DB 2; Length 191;

Pred. No. 2.5e-90;

Sequence 191 AA;

Query Match

Best Local Similarity

Score 1045; DB 2; Length 191;

Pred. No. 2.5e-90;

Sequence 191 AA;

Query Match

Best Local Similarity

Score 1045; DB 2; Length 191;

Pred. No. 2.5e-90;

Sequence 191 AA;

Query Match

Best Local Similarity

Score 1045; DB 2; Length 191;

Pred. No. 2.5e-90;

Sequence 191 AA;

Query Match

Best Local Similarity

Score 1045; DB 2; Length 191;

Pred. No. 2.5e-90;

Sequence 191 AA;

Query Match

Best Local Similarity

Score 1045; DB 2; Length 191;

Pred. No. 2.5e-90;

Sequence 191 AA;

Query Match

Best Local Similarity

Score 1045; DB 2; Length 191;

Pred. No. 2.5e-90;

Sequence 191 AA;

	Matches	191;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MSTLPEQRKTKRNTNRPTDVKFPFGGQIVGGVYLLPRRGPRGLGVATRTKTSRSQPRG	60							
Db	1	MSTLPEQRKTKRNTNRPTDVKFPFGGQIVGGVYLLPRRGPRGLGVATRTKTSRSQPRG	60							
Qy	61	RROQIPKARQPGQHWAQPYPWPPLYNGEGCGWALLSPRGSRPHWGPNDRPRSRLG	120							
Db	61	RROQIPKARQPGQHWAQPYPWPPLYNGEGCGWALLSPRGSRPHWGPNDRPRSRLG	120							
Qy	121	KVIDTLTCGFADLMGYIPVVGAPLGGAALAHGVRAIEDGINVATGNLPCGSFSIFLLA	180							
Db	121	KVIDTLTCGFADLMGYIPVVGAPLGGAALAHGVRAIEDGINVATGNLPCGSFSIFLLA	180							
Qy	181	LLSCLTTPPASA	191							
Db	181	LLSCLTTPPASA	191							

```

RESULT 2
AAR92972
ID AAR92972 standard; protein; 191 AA.
XX
AC AC
XX
XX 02-OCT-1996 (first entry)
XX
DE Hepatitis C virus isolate Z4 core protein.
DE DE
XX HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
KW hepatitis.
XX
XX Hepatitis C virus.
XX OS
XX WO9605315-A2.
XX PN
XX 22-FEB-1996.
XX PD
XX XX
XX PF 15-AUG-1995; 95WO-US010398.
XX XX
XX PR 15-AUG-1994; 94US-00290665.
XX XX
XX (USSH ) US SEC DEPT HEALTH.
XX PA
XX PI
XX Bukh J, Miller RH, Purcell RH;
XX WPI; 1996-139709/14.
XX DR N-PSDB: AAT16646.
XX DR

```

Qy	61	R R Q I P K A R Q Q G R H W A Q P G Y P W L Y G N E G C G W A L L S P R G S R P H W G P N D P R R S R N I G	120
Dd	61	R R Q I P K A R Q P E G S W A Q P G Y P W L Y G N E G C G W A L L S P R G S R P S W G P N D P R R S R N I G	120
Qy	121	K V I D T L T C G F A D L M G Y I P V V G A P L G G V A A L A H G V R A J E D G I N Y A T G N L P G C S F S I F L I A	180
Dd	121	K V I D T L T C G F A D L M G Y I P I V G A P V G G V A R A L A H G V R A V E D G I N Y A T G N L P G C S F S I F L I A	180
Qy	181	L L S C L T T P S A	191
Dd	181	L L S C L T T P S A	191
RESULT	3		
ID	AAR92974		
XX	AAR92974	standard; protein; 191 AA.	
AC	AAR92974;		

RESULT 3
AAR92974
ID AAR92974 standard; protein; 191 AA.
XX
XX AAR92974;
XX
XX
XX 02-OCT-1996 (first entry)
XX
XX Hepatitis C virus isolate Z1 core protein.
XX
XX HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
XX hepatitis.
XX
XX Hepatitis C virus.
XX OS
XX
XX WO9605315-A2.
XX
XX
XX 22-FEB-1996.
XX
XX
XX 15-AUG-1995; 95WO-US010398.
XX
XX 15-AUG-1994; 94US-00290665.
XX
XX (USSH) US SEC DEPT HEALTH.
XX
XX
XX Bukh J, Miller RH, Purcell RH;
XX
XX WPI; 1996-139709/14.
XX N-PSDB; AAT16648.
XX
XX DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
XX determine HCV genotype and as vaccines against HCV infection.
XX
XX Claim 4; Page 212; 340pp; English.

Query Match	94.0%	Score 982;	DB 2;	Length 191;
Best Local Similarity	93.7%;	Pred. No. 2.1e-84;		
Matches 179;	Conservative 5;	Mismatches 7;	Indels 0;	Gaps 0
Qy	1	MSTLPKQRTKRTNRRPTDVKPFPGGQIVGGVYLLPRRGPRLGVRATKRTSRSQPG	60	
Db	1	MSTNPKQRTKRTNRRPMDVKPFPGGQIVGGVYLLPRRGPRLGVRARKRTSRSQPG	60	
Qy	61	RRQIPKARQPGHHWAPQVPWPLYNNEGCGAGWLLSPGRSRPHWGPNDRPRSRNIG	120	
Db	61	RRQIPKARSEGSRWAPQVPWPLYNNEGCGAGWLLSPGRSRPSWGPNDRPRSRNIG	120	
Qy	121	KVIDTLTCGGADLMGYTPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFIFLLA	180	

```
Db 121 KVIDTLTCGFADLMGYIPLVGAPVGGVARALAHGVRVEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db |||||
181 LLSCLTTPASA 191

RESULT 4
AAR92973
ID AAR92973 standard; protein; 191 AA.
XX
AC
XX
DT 02-OCT-1996 (first entry)
XX
DE Hepatitis C virus isolate 28 core protein.
XX
KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
KW hepatitis.
XX
OS Hepatitis C virus.
XX
PN WO9605315-A2.
XX
PD 22-FEB-1996.
XX
PF 15-AUG-1995; 95WO-US010398.
XX
PR 15-AUG-1994; 94US-00290665.
XX
PA (USSH ) US SEC DEPT HEALTH.
PI Bukh J, Miller RH, Purcell RH;
XX
DR WPI; 1996-139709/14.
DR N-PSDB; AAT16647.
XX
PT DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
PT determine HCV genotype and as vaccines against HCV infection.
XX
PS Claim 4; Page 211-212; 340pp; English.
CC AAR92936-R92987 are HCV core proteins derived from 52 different HCV
CC isolates. Isolated cDNA sequences are used for the prodn. of primers
CC useful for detecting the presence of HCV in a sample, the primers are
CC also useful for HCV genotyping. Proteins encoded by the cDNAs can be used
CC in vaccines for immunising against HCV infection. The proteins may also
CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or
CC other mononuclear cells. The antibodies may be used in the prevention of
CC HCV infection
XX
SQ Sequence 191 AA;

Query Match 94.0%; Score 982; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 2.1e-84;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSTLPKQPKTKRNTNRRPTDVKFPGGGQIVGGVYLLPRGRPLGVRAATKTSERSQPRG 60
Db |||||
1 MSTNPKPQKTKRNTNRRPDMVKFPGGGQIVGGVYLLPRGRPLGVRAATKTSERSQPRG 60

QY 61 RROPKPKARQPOGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
Db |||||
1 MSTNPKPQKTKRNTNRRPDMVKFPGGGQIVGGVYLLPRGRPLGVRAATKTSERSQPRG 60

QY 61 RROPKPKARQPOGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
Db |||||
61 RROPKPKARQPOGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPLVGAPVGGVARALAHGVRVEDGINYATGNLPGCSFSIFLLA 180
Db |||||
121 KVIDTLTCGFADLMGYIPLVGAPVGGVARALAHGVRVEDGINYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
Db |||||
181 LLSCLTTPASA 191

RESULT 6
AAR92978
ID AAR92978 standard; protein; 191 AA.
XX
AC AAR92978;
```

```
RESULT 5
AAR92953
ID AAR92953 standard; protein; 191 AA.
XX
AC AAR92953;
XX
DT 02-OCT-1996 (first entry)
XX
DE Hepatitis C virus isolate HK3 core protein.
XX
KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
KW hepatitis.
XX
OS Hepatitis C virus.
XX
PN WO9605315-A2.
XX
PD 22-FEB-1996.
XX
PF 15-AUG-1995; 95WO-US010398.
XX
PR 15-AUG-1994; 94US-00290665.
XX
PA (USSH ) US SEC DEPT HEALTH.
PI Bukh J, Miller RH, Purcell RH;
XX
DR WPI; 1996-139709/14.
DR N-PSDB; AAT16627.
XX
PT DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
PT determine HCV genotype and as vaccines against HCV infection.
XX
PS Claim 4; Page 194-195; 340pp; English.
CC AAR92936-R92987 are HCV core proteins derived from 52 different HCV
CC isolates. Isolated cDNA sequences are used for the prodn. of primers
CC useful for detecting the presence of HCV in a sample, the primers are
CC also useful for HCV genotyping. Proteins encoded by the cDNAs can be used
CC in vaccines for immunising against HCV infection. The proteins may also
CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or
CC other mononuclear cells. The antibodies may be used in the prevention of
CC HCV infection
XX
SQ Sequence 191 AA;

Query Match 93.9%; Score 981; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 2.6e-84;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSTLPKQPKTKRNTNRRPTDVKFPGGGQIVGGVYLLPRGRPLGVRAATKTSERSQPRG 60
Db |||||
1 MSTNPKPQKTKRNTNRRPQDVKFPFGGGQIVGGVYLLPRGRPLGVRAATKTSERSQPRG 60

QY 61 RROPKPKARQPOGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
Db |||||
61 RROPKPKARQPOGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPLVGAPVGGVARALAHGVRVEDGINYATGNLPGCSFSIFLLA 180
Db |||||
121 KVIDTLTCGFADLMGYIPLVGAPVGGVARALAHGVRVEDGINYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
Db |||||
181 LLSCLTTPASA 191

RESULT 6
AAR92978
ID AAR92978 standard; protein; 191 AA.
XX
AC AAR92978;
```

XX 02-OCT-1996 (first entry)
 DT Hepatitis C virus isolate DK13 core protein.
 DE HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
 XX hepatitis.
 KW Hepatitis C virus.
 OS WO9605315-A2.
 XX 22-FEB-1996.
 PD
 XX Hepatitis C virus.
 OS
 XX WO9605315-A2.
 PN
 XX 22-FEB-1996.
 PD
 XX 15-AUG-1995; 95WO-US010398.
 PF
 XX 15-AUG-1994; 94US-00290665.
 PR
 XX (USSH) US SEC DEPT HEALTH.
 PA
 XX Bukh J, Miller RH, Purcell RH;
 PI
 XX WPI; 1996-139709/14.
 DR
 XX N-PSDB; AAT16652.
 DR
 XX DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
 PT determine HCV genotype and as vaccines against HCV infection.
 PT
 XX Claim 4; Page 215-216; 340pp; English.
 PS
 XX AAR92936-R92987 are HCV core proteins derived from 52 different HCV
 CC isolates. Isolated cDNA sequences are used for the prodn. of primers
 CC useful for detecting the presence of HCV in a sample, the primers are
 CC also useful for HCV genotyping. Proteins encoded by the cDNAs can be used
 CC in vaccines for immunising against HCV infection. The proteins may also
 CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or
 CC other mononuclear cells. The antibodies may be used in the prevention of
 CC HCV infection
 CC
 XX Sequence 191 AA;
 SQ
 Query Match 93.8%; Score 980; DB 2; Length 191;
 Best Local Similarity 93.7%; Pred. No. 3.3e-84;
 Matches 179; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 Qy 1 MSTILPKPQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVATRTKTSERSQPRG 60
 Db 1 MSTNPKPQRTKNTNRPMDDVKFPGGGQIVGGVYLLPRGPRGLGVATRTKTSERSQPRG 60
 Qy 61 RRQIPKARQPOGRHWAQPGYPWPLYNCGCGWAGWLLSPRSGRPHWGPNDPRRSRNLG 120
 Db 61 RRQIPKARQLEGRSWAQPGYPWPLYNCGCGWAGWLLSPRSGRPSWGPNDPRRSRNLG 120
 Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
 Db 121 KVIDTLTCGFADLMGYIPVVGAPVGGVARALAHGVRLLIEDGVNATGNLPGCSFSIFLLA 180
 Qy 181 LLSCLTTPASA 191
 Db 181 LLSCLTVPASA 191
 RESULT 7
 AAR92976
 ID AAR92976 standard; protein; 191 AA.
 AC
 XX AAR92976;
 AC
 XX 02-OCT-1996 (first entry)
 DT
 XX Hepatitis C virus isolate Z6 core protein.
 DE
 XX HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
 KW hepatitis.

hepatitis.
 KW
 XX Hepatitis C virus.
 OS
 XX WO9605315-A2.
 PN
 XX 22-FEB-1996.
 PD
 XX 15-AUG-1995; 95WO-US010398.
 PF
 XX 15-AUG-1994; 94US-00290665.
 PR
 XX (USSH) US SEC DEPT HEALTH.
 PA
 XX Bukh J, Miller RH, Purcell RH;
 PI
 XX WPI; 1996-139709/14.
 DR
 XX N-PSDB; AAT16650.
 DR
 XX DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
 PT determine HCV genotype and as vaccines against HCV infection.
 PT
 XX Claim 4; Page 214; 340pp; English.
 PS
 XX AAR92936-R92987 are HCV core proteins derived from 52 different HCV
 CC isolates. Isolated cDNA sequences are used for the prodn. of primers
 CC useful for detecting the presence of HCV in a sample, the primers are
 CC also useful for HCV genotyping. Proteins encoded by the cDNAs can be used
 CC in vaccines for immunising against HCV infection. The proteins may also
 CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or
 CC other mononuclear cells. The antibodies may be used in the prevention of
 CC HCV infection
 CC
 XX Sequence 191 AA;
 SQ
 Query Match 93.6%; Score 978; DB 2; Length 191;
 Best Local Similarity 93.2%; Pred. No. 5.1e-84;
 Matches 178; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 Qy 1 MSTILPKPQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVATRTKTSERSQPRG 60
 Db 1 MSTNPKPQRTKNTNRPMDDVKFPGGGQIVGGVYLLPRGPRGLGVATRTKTSERSQPRG 60
 Qy 61 RRQIPKARQPOGRHWAQPGYPWPLYNCGCGWAGWLLSPRSGRPHWGPNDPRRSRNLG 120
 Db 61 RRQIPKARSEGSAQPGYPWPLYNCGCGWAGWLLSPRSGRPSWGPNDPRRSRNLG 120
 Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
 Db 121 KVIDTLTCGFADLMGYIPVVGAPVGGVARALAHGVRVEDGINVATGNLPGCSFSIFLLA 180
 Qy 181 LLSCLTTPASA 191
 Db 181 LLSCLTVPASA 191
 RESULT 8
 AAR92977
 ID AAR92977 standard; protein; 191 AA.
 AC
 XX AAR92977;
 AC
 XX 02-OCT-1996 (first entry)
 DT
 XX Hepatitis C virus isolate Z7 core protein.
 DE
 XX HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
 KW hepatitis.
 KW
 XX Hepatitis C virus.
 OS
 XX WO9605315-A2.
 PN
 XX

PD 22-FEB-1996.
 XX 15-AUG-1995; 95WO-US010398.
 XX 15-AUG-1994; 94US-00290665.
 XX (USSH) US SEC DEPT HEALTH.
 XX Bukh J, Miller RH, Purcell RH;
 XX WPI; 1996-139709/14.
 DR N-PSDB; AAT16651.
 XX DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
 PT determine HCV genotype and as vaccines against HCV infection.
 XX Claim 4; Page 214-215; 340pp; English.
 XX AAR92936-R92987 are HCV core proteins derived from 52 different HCV
 CC isolates. Isolated cDNA sequences are used for the prodn. of primers
 CC useful for detecting the presence of HCV in a sample, the primers are
 CC also useful for HCV genotyping. Proteins encoded by the cDNAs can be used
 CC in vaccines for immunising against HCV infection. The proteins may also
 CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or
 CC other mononuclear cells. The antibodies may be used in the prevention of
 CC HCV infection
 XX Sequence 191 AA;
 SQ Query Match 93.5%; Score 977; DB 2; Length 191;
 Best Local Similarity 93.2%; Pred. No. 6.3e-84;
 Matches 178; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MSTLPKQKTKRNTNRRPTDVKFPGGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
 DB 1 MSTNPKPQKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
 QY 61 RQPIPKARQPOGRHWAQPGYPMPLYGNECGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
 DB 61 RQPIPKARSRGRSQAQPGYPMPLYGNECGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
 QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVAIEDGINYATGNLPGCSFSIFLLA 180
 DB 121 KVIDTLTCGFADLMGYIPLVGAPVGVARALAHGVALEDDGINYATGNLPGCSFSIFLLA 180
 QY 181 LLSCLTTPASA 191
 DB 181 LLSCLTVPASA 191
 RESULT 9
 AAR74048
 ID AAR74048 standard; protein; 196 AA.
 AC AAR74048;
 XX 25-MAR-2003 (revised)
 DT 26-NOV-1995 (first entry)
 XX Synthetic HCV nucleocapsid protein.
 XX Hepatitis C virus; vector; expression; prokaryotic cells; detection;
 KW diagnosis.
 XX Synthetic.
 OS WO9511980-A2.
 XX 04-MAY-1995.
 XX 25-OCT-1994; 94WO-US012166.
 XX 25-OCT-1993; 93US-00141917.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Khudyakov Y, Fields HA;
 XX WPI; 1995-178872/23.
 XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Khudyakov Y, Fields HA;
 XX WPI; 1995-178872/23.
 DR N-PSDB; AAQ92071.
 XX Vector for expression of a synthetically produced protein coding DNA -
 PT pref. encodes the hepatitis C virus nucleocapsid protein which can be
 PT used in the detection of HCV antibodies.
 XX Disclosure; Page 37; 44pp; English.
 XX The sequence is that of a synthetic nucleotide sequence encoding the
 CC hepatitis C virus nucleocapsid protein. The gene is positioned in a
 CC vector for efficient expression in prokaryotic cells. The expressed
 CC protein can be used in tests for the detection of antibodies specific for
 CC the HCV C protein. See also AAR74047. (Updated on 25-MAR-2003 to correct
 CC PN field.)
 XX Sequence 196 AA;
 SQ Query Match 93.5%; Score 977; DB 2; Length 196;
 Best Local Similarity 92.1%; Pred. No. 6.5e-84;
 Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
 QY 1 MSTLPKQKTKRNTNRRPTDVKFPGGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
 DB 1 MSTIPKQKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
 QY 61 RQPIPKARQPOGRHWAQPGYPMPLYGNECGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
 DB 61 RQPIPKVRPRPGRTWAQPGYPMPLYGNECGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
 QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVAIEDGINYATGNLPGCSFSIFLLA 180
 DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGARALAHGVRVEDGVNYATGNLPGCSFSIFLLA 180
 QY 181 LLSCLTTPASA 191
 DB 181 LLSCLTVPASA 191
 RESULT 10
 AAR74047
 ID AAR74047 standard; protein; 196 AA.
 AC AAR74047;
 XX 25-MAR-2003 (revised)
 DT 26-NOV-1995 (first entry)
 XX Synthetic HCV nucleocapsid protein.
 XX Hepatitis C virus; vector; expression; prokaryotic cells; detection;
 KW diagnosis.
 XX Synthetic.
 OS WO9511980-A2.
 XX 04-MAY-1995.
 XX 25-OCT-1994; 94WO-US012166.
 XX 25-OCT-1993; 93US-00141917.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Khudyakov Y, Fields HA;
 XX WPI; 1995-178872/23.
 XX

DR N-PSDB; AAQ92070.

XX Vector for expression of a synthetically produced protein coding DNA -

PT pref. encodes the hepatitis C virus nucleocapsid protein which can be

PT used in the detection of HCV antibodies.

XX

XX Disclosure; Page 35; 44pp; English.

XX

CC The sequence is that of a synthetic nucleotide sequence encoding the

CC hepatitis C virus nucleocapsid protein. The gene is positioned in a

CC vector for efficient expression in prokaryotic cells. The expressed

CC protein can be used in tests for the detection of antibodies specific for

CC the HCV C protein. See also AAR74048. (Updated on 25-MAR-2003 to correct

CC PN field.)

XX

SQ Sequence 196 AA;

Query Match 93.5%; Score 977; DB 2; Length 196;

Best Local Similarity 92.1%; Pred. No. 6.5e-84;

Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MSTLPKQRTKRTNRRPTDVKFPGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60

Db 1 MSTLPKQRTKRTNRRPTDVKFPGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60

Qy 61 RRQPIPKAROPQGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGPNDRSRNMG 120

Db 61 RRQPIPKAROPQGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGPNDRSRNMG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVATGNLPCGCSFIFLLA 180

Db 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVATGNLPCGCSFIFLLA 180

Qy 181 LLSCLTTPASA 191

Db 181 LLSCLTTPASA 191

RESULT 11

AAR96546

ID AAR96546 standard; peptide; 319 AA.

XX AC AAR96546;

XX AC

XX 10-MAR-1997 (first entry)

XX Hepatitis C virus types 9a(7a) isolates FR1 amino acids 1-317.

XX Hepatitis C virus; subtype; polymerase chain reaction; amplification;

KW PCR; primer; probe; antibody; infection.

XX Hepatitis C virus.

OS

Key Location/Qualifiers

FT Misc-difference 128

FT /label= Phe, Ser, Tyr, Cys

FT Misc-difference 129

FT /label= Gly

FT /note= "amino acid in this position is designated X in

FT the specification, but codon usage shows that the only

FT possible amino acid at this pos. is Gly"

FT Misc-difference 308

FT /label= Ile, Met

FT Misc-difference 310

FT /label= Ser, Pro, Thr, Ala

XX

PN WO9613590-A2.

XX

PD 09-MAY-1996.

XX

PF 23-OCT-1995; 95WO-EP004155.

XX

PR 21-OCT-1994; 94EP-00870166.

PR 28-JUN-1995; 95EP-00870076.

XX (INNO-) INNOGENETICS NV.

PA Maertens G, Stuyver L;

XX WPI; 1996-251460/25.

DR N-PSDB; AAT27957.

XX

XX Hepatitis C virus polynucleic acid unique to unidentified sub-type -

PT used to develop probes and primers for new subtypes and vaccines to

PT prevent and treat infection.

XX

XX Claim 25; Fig 3; 150pp; English.

XX

CC The sequences AAR96526-R96578 represent novel sequences isolated from

CC hepatitis C virus subtypes different from subtypes 1a-c, 2a-d, 3a-f, 4a-

CC j, 5a and 6a. They esp. from the novel subtypes id-f, 2e-1, 2k, 2l, 3g,

CC 4k-m, 7a-c or types 9, 10 or 11. The sequences corresp. to the 5'

CC untranslated region (UR), the Core/E1, NS4 or NS5B regions of the genome.

CC This sequence represents amino acids 1-317 from the HCV types 9a and 7a

CC isolates FR1. The new HCV types were isolated from patients with chronic

CC HCV from the Benelux countries, France, Cameroon and Vietnam, because of

CC their aberrant reactivities. The RNA was extracted, cDNA synthesised and

CC PCR amplified, cloned and genotyped. The 5'UR, Core/E1 and NS5B regions

CC were sequenced either directly or partially and used to classify the new

CC viruses into (sub)types based on comparison with known sequences. The

CC sequences were also used to generate the peptides AAR96424-R96524. The

CC sequences can also be used to synthesise probes and primers for the

CC detection of HCV in a sample. The polypeptides can be used to detect anti

CC -HCV antibodies, for HCV typing or to prevent HCV infections

XX

SQ Sequence 319 AA;

Query Match 93.5%; Score 977; DB 2; Length 319;

Best Local Similarity 93.7%; Pred. No. 1.1e-83;

Matches 179; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MSTLPKQRTKRTNRRPTDVKFPGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60

Db 1 MSTLPKQRTKRTNRRPTDVKFPGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60

Qy 61 RRQPIPKAROPQGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGPNDRSRNMG 120

Db 61 RRQPIPKAROPQGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGPNDRSRNMG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVATGNLPCGCSFIFLLA 180

Db 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVATGNLPCGCSFIFLLA 180

Qy 181 LLSCLTTPASA 191

Db 181 LLSCLTTPASA 191

RESULT 12

AAR22137

ID AAR22137 standard; protein; 326 AA.

XX AC AAR22137;

XX 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 07-JUL-1992 (first entry)

XX HCV-Hc59 capsid and envelope proteins.

XX Hepatitis C virus; non-A non-B virus; HCV-Hc59; antigen; vaccine; assay;

KW detection.

XX Non-A.

OS non-B hepatitis virus.

XX

Db 181 LLSCLTVPASA 191

RESULT 14
AAR24440
ID AAR24440 standard; protein; 2894 AA.
AC AAR24440;
XX
XX 25-MAR-2003 (revised)
DT 02-DEC-1992 (first entry)
XX
XX Composite HCV HC-J1/CDC/CHI protein.
XX Hepatitis C virus; peptides; antibodies; ELISA.
XX Synthetic.

Key Location/Qualifiers
FT Peptide 1..20
FT Peptide /label= 1
FT Peptide 7..26
FT Peptide /label= 2
FT Peptide 13..32
FT Peptide /label= 3
FT Peptide 37..56
FT Peptide /label= 4
FT Peptide 49..68
FT Peptide /label= 5
FT Peptide 61..80
FT Peptide /label= 6
FT Peptide 73..92
FT Peptide /label= 7
FT Peptide 168..1707
FT Peptide /label= 8
FT Peptide 1694..1713
FT Peptide /label= 9
FT Peptide 1706..1725
FT Peptide /label= 10
FT Peptide 1712..1731
FT Peptide /label= 11
FT Peptide 1718..1737
FT Peptide /label= 12
FT Peptide 1724..1743
FT Peptide /label= 13
FT Peptide 1730..1749
FT Peptide /label= 14
FT Peptide 2263..2282
FT Peptide /label= 15
FT Peptide 2275..2294
FT Peptide /label= 16
FT Peptide 2287..2306
FT Peptide /label= 17
FT Peptide 2299..2318
FT Peptide /label= 18
FT Peptide 2311..2330
FT Peptide /label= 19
XX
XX EP489968-A1.
XX
XX 17-JUN-1992.
XX
XX 14-DEC-1990; 90EP-00124241.
XX
XX 14-DEC-1990; 90EP-00124241.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Deleys RJ, Pollet D, Maertens G, Van Heuverswyn H;
XX
XX WPI; 1992-201383/25.
XX
XX New synthetic peptides for detecting antibodies to hepatitis C virus -

PT useful in e.g. ELISA assays, and for detection of HCV antigens or as immunogens.
XX Disclosure; Fig 1; 32pp; English.
XX RNA viruses frequently exhibit a high rate of spontaneous mutation, thus a virus is considered to be the same of equiv. to HCV if it exhibits a global homology of more than 70 percent with the HCV HC-J1/CDC/CHI composite sequence. The peptide fragments of this DNA sequence indicated in the features table can immunologically mimic proteins encoded by HCV.
CC Additional amino acids or chemical gps. may be added to either end of the peptides for the purpose of creating a linker arm for attachment to a carrier. The peptides can be used for the detection of antibodies specific for HCV. They may be used in the form of kits, opt. with reagents such as staphylococcal protein A, streptococcal protein G, avidin or streptavidin. The peptides may also be used as immunogens for raising antibodies. (Updated on 25-MAR-2003 to correct FN field.)
XX
XX Sequence 2894 AA;
SQ

Query Match 93.5%; Score 977; DB 2; Length 2894;
Best Local Similarity 92.1%; Pred. No. 1.3e-82;
Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MSTLPKPKRKTNRNRTDVKFPGGGQIVGGVYLLPRRGLVATRTKTSRSQPRG 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MSTIPKPKRKTNRNRRPDVKFPGGGQIVGGVYLLPRRGLVATRTKTSRSQPRG 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 61 RRPIPKAROPQGRHWAQPGYWPPLYNCGCGWAGWLLSPRGSPHWPNDPPRSNIG 120
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 RRQPIPKVREPEGRWTWQPGYWPPLYNCGCGWAGWLLSPRGSPHWPNDPPRSNIG 120
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 181 LLSCLTVPASA 191
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 LLSCLTVPASA 191
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 15
AAR70230
ID AAR70230 standard; protein; 2894 AA.
XX
XX AAR70230;
AC AAR70230;
XX
XX 25-MAR-2003 (revised)
DT 07-NOV-1995 (first entry)
XX
XX Composite hepatitis C virus (HC-J1/CDC/CHI).
XX
XX Composite hepatitis C virus; HC-J1/CDC/CHI; HCV; non-A non-B;
XX synthetic antigens; blood screening.
XX
XX Hepatitis C virus.
XX
XX EP644202-A1.
XX
XX 22-MAR-1995.
XX
XX 14-DEC-1990; 94EP-00108611.
XX
XX 14-DEC-1990; 90EP-00124241.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Deleys RJ, Pollet D, Maertens G, Van Heuverswyn H;
XX
XX WPI; 1995-116946/16.
XX
XX Synthetic antigens for the detection of hepatitis C virus antibodies -
PT comprise portions of the HCV peptide sequence, for use in screening blood

PT and blood products.
XX
PS Disclosure; Fig 1; 51pp; English.
XX
CC AAR70230 is the composite hepatitis C virus (HC-J1/CDC/CHI) protein from
CC which the synthetic HCV antigens described in AAR70210-R70229 were
CC derived. These synthetic antigens can be used to screen blood, or blood
CC products for the presence of HCV, they can also be used in various specific
CC assays for the detection of HCV antibodies, and antigens, or as
CC immunogens. (Updated on 25-MAR-2003 to correct FN field.) (Updated on 25-
CC MAR-2003 to correct PF field.)
XX
SQ Sequence 2894 AA;

Query Match 93.5%; Score 977; DB 2; Length 2894;
Best Local Similarity 92.1%; Pred. No. 1.3e-82;
Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
QY 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVLLPRRGPRLGVRATKTSERSQPRG 60
|||:|||||
Db 1 MSTLPKQKTKRNTNRRPQDVKFGGGQIVGGVLLPRRGPRLGVRATKTSERSQPRG 60
|||:|||||
QY 61 RRQPIPKARPOGQRHWAQGYWPPLYGNEGCGWAGWLLSPRGSRPHWGPNDRSRNLG 120
|||:|||||
Db 61 RRQPIPKVRRPEGRRTWAQGYWPPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120
|||:|||||
QY 121 KVDTLTTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
|||:|||||
Db 121 KVDTLTTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
|||:|||||
QY 181 LLSCLTTPASA 191
|||:|||||
Db 181 LLSCLTVPASA 191
|||:|||||

Search completed: October 30, 2004, 01:42:44
Job time : 157 secs

QY 181 LLSCLTTPASA 191
Db 181 LLSCLTIPASA 191

RESULT 5
PN0677
hypotheoretical protein 787 - hepatitis C virus (fragment)
C:Species: hepatitis C virus
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C:Accession: PN0677
R:Cho, S.H.; Yoon, J.I.; Chang, J.E.; Ahn, B.M.; Lee, C.H.; Lee, Y.I.
Biochem. Biophys. Res. Commun. 196; 780-789, 1993
A:Title: Genomic typing of hepatitis C viruses from Korean patients: Implications of gen
A:Reference number: PN0677; PMID:94059104; PMID:8240354
A:Accession: PN0677
A:Molecule type: mRNA
A:Residues: 1-787 <CHO>
A:Cross-references: UNIPROT:O08244; GB:L20498; NID:gl381031; PIDN:AA02608.1; PID:gl3810
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: glycoprotein; nonstructural protein
F:196,209,234,250,305,325,421,427,452,536,544,560,580,627,649/Binding site: carbohydrate
Query Match 92.7%; Score 969; DB 2; Length 787;
Best Local Similarity 92.7%; Pred. No. 3.8e-74;
Matches 177; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 MSTLPKPKRKNTNRPTDVKPGGQIVGGVYLLPRGRPLGVRATKTSERSQPRG 60
Db 1 MSTNPKPKRKNTNRPPQDVKPGGQIVGGVYLLPRGRPLGVRATKTSERSQPRG 60

QY 61 RROPTPKARQPGRWAOQGYWPPLYGNEGCWAGWLLSPGSRPHWGPNDRSRNLG 120
Db 61 RROPTPKARQPGRWAOQGYWPPLYGNEGCWAGWLLSPGSRPHWGPNDRSRNLG 120

QY 121 KVIDLTGCFADLMGYIPVWGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDLTGCFADLMGYIPLVGNAPLGGARALAHGVRLVDGYNVATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
Db 181 LLSCLTIPASA 191

RESULT 6
GNWVCJ
genome polypeptide - hepatitis C virus (strain HCV-1)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polypeptide
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C:Accession: A39166; P04043
R:Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Co
Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A:Title: Genetic organization and diversity of the hepatitis C virus.
A:Reference number: A39166; PMID:91172826; PMID:1848704
A:Accession: A39166
A:Molecule type: mRNA
A:Residues: 1-3011 <CHO>
A:Cross-references: UNIPROT:P26664; GB:M62321; NID:G329873; PIDN:AAA45676.1; PID:G329874
R:Chan, S.W.; McMahon, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.I
J. Gen. Virol. 73, 1131-1141, 1992
A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A:Reference number: P04043; PMID:92268871; PMID:1316939
A:Accession: P04043
A:Molecule type: genomic RNA
A:Residues: 1577-1633 <CHA>
A:Cross-references: DDBJ:D10128
A:Experimental source: isolates E-b16
A:Accession: P0404
A:Status: preliminary
A:Molecule type: genomic RNA

A:Residues: 1577-1633 <CH2>
A:Experimental source: isolates E-b17
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F:1115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEB>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polypeptide <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077,22

Query Match 92.6%; Score 968; DB 1; Length 3011;
Best Local Similarity 91.6%; Pred. No. 1.6e-73;
Matches 175; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 MSTLPKPKRKNTNRPTDVKPGGQIVGGVYLLPRGRPLGVRATKTSERSQPRG 60
Db 1 MSTNPKPKRKNTNRPPQDVKPGGQIVGGVYLLPRGRPLGVRATKTSERSQPRG 60

QY 61 RROPTPKARQPGRWAOQGYWPPLYGNEGCWAGWLLSPGSRPHWGPNDRSRNLG 120
Db 61 RROPTPKARQPGRWAOQGYWPPLYGNEGCWAGWLLSPGSRPHWGPNDRSRNLG 120

QY 121 KVIDLTGCFADLMGYIPVWGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDLTGCFADLMGYIPLVGNAPLGGARALAHGVRLVDGYNVATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
Db 181 LLSCLTIPASA 191

RESULT 7
GNWVCJ
genome polypeptide - hepatitis C virus (strain J)
N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructura
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A39253; PS0086
R:Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimot
Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990
A:Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients
A:Reference number: A39253; PMID:91088550; PMID:2175903
A:Accession: A39253
A:Molecule type: genomic RNA
A:Residues: 1-3010 <KAT>
A:Cross-references: UNIPROT:P26662; GB:D90208; NID:G221610; PIDN:BAAL4233.1; PID:G221611
R:Kato, N.; Ohkoshi, S.; Shimotohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989
A:Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence vari
A:Reference number: PS0085
A:Accession: PS0086
A:Molecule type: genomic RNA
A:Residues: 2650-2707 <KAT>
A:Experimental source: Japanese isolate
C:Comment: The cleavage sites of this polypeptide have not been determined.
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEB>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polypeptide <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)

F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif

F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4b>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,2

Query Match 92.3%; Score 965; DB 1; Length 3010;
Best Local Similarity 92.1%; Pred. No. 2.9e-73;
Matches 176; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MSTLPKQRTKNTNRPTDVKPPGGQIVGGVYLLPRRGLVGRVATRKTSRSOPRG 60

Db 1 MSTNPKQRTKNTNRPPQDVKPPGGQIVGGVYLLPRRGLVGRVATRKTSRSOPRG 60

Qy 61 RRQIPKARPOQRHWAQPGYPWPLYNCGGAGWLLSPRGSPHGWPNDRPRRSNRLG 120

Db 61 RRQIPKARPEGRWTAPQGPWPLYNCGGAGWLLSPRGSPHGWPNDRPRRSNRLG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPCSCSIFLLA 180

Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGVAAALAHGVRAIEDGINVATGNLPCSCSIFLLA 180

Qy 181 LLSCLTTPASA 191

Db 181 LLSCLTTPASA 191

RESULT 8

GNWYTC

genome polyprotein - hepatitis C virus
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein NS5
Protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C:Species: hepatitis C virus

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004

C:Accession: A38465

R:Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.;

A:Title: Structure and organization of the hepatitis C virus genome isolated from human

A:Reference number: A38465; MUID:91140698; PMID:1847440

A:Accession: A38465

A:Molecule type: genomic RNA

A:Residues: 1-3010 <TAK>

A:Cross-references: UNIPROT:P26663; EMBL:M58335; NID:G329770; PIDN:AAA72945.1; PID:G3297

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural

F:2-115/Product: capsid protein C #status predicted <CPC>

F:116-191/Product: envelope protein M #status predicted <EPM>

F:192-389/Product: major envelope protein E #status predicted <MEE>

F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

F:1007-1615/Product: hepatitis C virus genome polyprotein NS3 #status predicted <NS3>

F:1230-1237/Region: nucleotide-binding motif A (P-loop)

F:1312-1317/Region: nucleotide-binding motif B

F:1316-1319/Region: DEXH motif

F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>

F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4b>

F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

F:196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,22

Query Match 92.3%; Score 965; DB 1; Length 3010;

Best Local Similarity 92.1%; Pred. No. 2.9e-73;

Matches 176; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MSTLPKQRTKNTNRPTDVKPPGGQIVGGVYLLPRRGLVGRVATRKTSRSOPRG 60

Db 1 MSTNPKQRTKNTNRPPQDVKPPGGQIVGGVYLLPRRGLVGRVATRKTSRSOPRG 60

Qy 61 RRQIPKARPOQRHWAQPGYPWPLYNCGGAGWLLSPRGSPHGWPNDRPRRSNRLG 120

Db 61 RRQIPKARPEGRWTAPQGPWPLYNCGGAGWLLSPRGSPHGWPNDRPRRSNRLG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPCSCSIFLLA 180

Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGVAAALAHGVRAIEDGINVATGNLPCSCSIFLLA 180

Qy 181 LLSCLTTPASA 191

Db 181 LLSCLTTPASA 191

RESULT 9

JH0711

genome polyprotein - hepatitis C virus (strain PRC1) (fragments)

N:Contains: envelope protein E1; envelope protein E2; nonstructural protein NS1; nonstru

C:Species: hepatitis C virus

C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004

C:Accession: JH0711

R:Liu, K.; Hu, Z.; Li, H.; Prince, A.M.; Inchauspe, G.

Gene 114, 245-250, 1992

A:Title: Genomic typing of hepatitis C viruses present in China.

A:Reference number: JH0711; MUID:92290283; PMID:1318245

A:Accession: JH0711

A:Molecule type: genomic RNA

A:Residues: 1-550 <LIU>

A:Cross-references: UNIPROT:Q7LZV4; GB:M74888; GB:M74889

A:Note: the nucleotide sequence is not complete

A:Note: translation of the nucleotide sequence is not complete

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: envelope protein; glycoprotein; nonstructural protein; nucleocapsid; polypro

F:1-190/Product: nucleocapsid protein C #status predicted <CPC>

F:191-380/Product: envelope protein E1 #status predicted <EPI>

F:381-514/Product: envelope protein E2 or nonstructural protein NS1 (fragment) #status

F:515-550/Product: nonstructural protein NS5 (fragment) #status predicted <NS5>

F:196,233,250,305,416,422,429,447/Binding site: carbohydrate (Asn) #status pr

Query Match 92.1%; Score 962; DB 2; Length 550;

Best Local Similarity 91.6%; Pred. No. 1.1e-73;

Matches 175; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MSTLPKQRTKNTNRPTDVKPPGGQIVGGVYLLPRRGLVGRVATRKTSRSOPRG 60

Db 1 MSTNPKQRTKNTNRPPQDVKPPGGQIVGGVYLLPRRGLVGRVATRKTSRSOPRG 60

Qy 61 RRQIPKARPOQRHWAQPGYPWPLYNCGGAGWLLSPRGSPHGWPNDRPRRSNRLG 120

Db 61 RRQIPKARPEGRWTAPQGPWPLYNCGGAGWLLSPRGSPHGWPNDRPRRSNRLG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPCSCSIFLLA 180

Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGVAAALAHGVRAIEDGINVATGNLPCSCSIFLLA 180

Qy 181 LLSCLTTPASA 191

Db 181 LLSCLTTPASA 191

RESULT 10

JQ1584

genome polyprotein - hepatitis C virus (strain U.K.) (fragment)

N:Contains: core protein C; envelope protein E1; envelope protein E2; nonstructural pro

C:Species: hepatitis C virus

C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004

C:Accession: JQ1584

R:Kumar, U.; Cheng, D.; Thomas, H.; Monjardino, J.

J. Gen. Virol. 73, 1521-1525, 1992

A:Title: Cloning and sequencing of the structural region and expression of putative cor

A:Reference number: JQ1584; MUID:92300349; PMID:1318944

A:Molecule type: genomic RNA

A:Residues: 1-640 <KUM>

A:Cross-references: UNIPROT:Q68966; GB:X84079; NID:G643119; PIDN:CAA58888.1; PID:G64312

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: core protein; envelope protein; glycoprotein; nonstructural protein; polypro

F:1-191/Product: core protein C #status predicted <CPC>

F:192-369/Product: envelope protein E1 #status predicted <EE1>

F:390-640/Product: envelope protein E2 and nonstructural protein NS1 #status predicted <F:196,209,234,305,417,430,448,476,540,556,576,623/Binding site: carbohydrate (cov

Query Match 92.0%; Score 961; DB 2; Length 640;
Best Local Similarity 91.1%; Pred. No. 1.5e-73;
Matches 174; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFPGGGQIVGGVLLPRRGLGVRAIRKTSERSQPRG 60
DB 1 MSTNPKPKQKTKRNTNRRPQDVKFGGGQIVGGVLLPRRGLGVRAIRKTSERSQPRG 60
QY 61 RROPIPKARQPOGRHWAQPGYWPVLYGNEGCGWAGLLSPGRSRPHWGNDRPRRSRLG 120
DB 61 RROPIPKARPEGRWAQPGYWPVLYGNEGCGWAGLLSPGRSRPHWGNDRPRRSRLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 11
S32740
polyprotein - hepatitis C virus (isolate Russian) (fragment)
N:Contains: capsid protein C; envelope protein M
C:Species: hepatitis C virus
C:Date: 19-Mar-1997 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
C:Accession: S32740

R:Vassilev, V.B.; Viazov, S.O.; Kotova, E.Y.; Nosikov, V.V.
submitted to the EMBL Data Library, April 1993
A:Description: Evidence of new HCV variant of European isolate in Russia.

A:Reference number: S32740
A:Accession: S32740
A:Molecule type: genomic RNA
A:Residues: 1-189 <VAS>
A:Cross-references: UNIPROT:Q68873; EMBL:X71407
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; envelope protein; polyprotein
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-189/Product: envelope protein M #status predicted <EPM>

Query Match 91.8%; Score 959; DB 2; Length 189;
Best Local Similarity 92.1%; Pred. No. 6.9e-74;
Matches 174; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFPGGGQIVGGVLLPRRGLGVRAIRKTSERSQPRG 60
DB 1 MSTNPKPKQKTKRNTNRRPQDVKFGGGQIVGGVLLPRRGLGVRAIRKTSERSQPRG 60
QY 61 RROPIPKARQPOGRHWAQPGYWPVLYGNEGCGWAGLLSPGRSRPHWGNDRPRRSRLG 120
DB 61 RROPIPKARPEGRWAQPGYWPVLYGNEGCGWAGLLSPGRSRPHWGNDRPRRSRLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPA 189
DB 181 LLSCLTTPA 189

RESULT 12
S19876
genome polyprotein - hepatitis C virus (isolate JK5) (fragment)
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
C:Species: hepatitis C virus
A:Variety: isolate JK5
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: S19876

R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.

submitted to the EMBL Data Library, September 1991
A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso

A:Reference number: S18029
A:Accession: S19876
A:Molecule type: genomic RNA
A:Residues: 1-782 <HON>

A:Cross-references: UNIPROT:Q68953; EMBL:X61595; NID:G59486; PIDN:CAA43792.1; PID:G59487
A:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
F:1-191/Product: core protein #status predicted <MAT1>
F:192-383/Product: envelope protein 1 #status predicted <MAT2>
F:384-733/Product: NS1/E2 protein #status predicted <MAT3>
F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match 91.7%; Score 958; DB 2; Length 782;
Best Local Similarity 91.6%; Pred. No. 3.2e-73;
Matches 175; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFPGGGQIVGGVLLPRRGLGVRAIRKTSERSQPRG 60
DB 1 MSTNPKPKQKTKRNTNRRPQDVKFGGGQIVGGVLLPRRGLGVRAIRKTSERSQPRG 60
QY 61 RROPIPKARQPOGRHWAQPGYWPVLYGNEGCGWAGLLSPGRSRPHWGNDRPRRSRLG 120
DB 61 RROPIPKARPEGRWAQPGYWPVLYGNEGCGWAGLLSPGRSRPHWGNDRPRRSRLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 13

PC1284
genome polyprotein - hepatitis C virus (isolate HC-J4) (fragment)
C:Species: hepatitis C virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PC1284
R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990
A:Title: The 5'-terminal sequence of the hepatitis C virus genome.
A:Reference number: PC1284; MUID:91013116; PMID:2170712
A:Accession: PC1284

A:Molecule type: genomic RNA
A:Residues: 1-513 <OK>
A:Cross-references: UNIPROT:Q81221; GB:D00832; NID:G221513; PIDN:BAA00706.1; PID:G221514
C:Superfamily: hepatitis C virus genome polyprotein

Query Match 91.6%; Score 957; DB 2; Length 513;
Best Local Similarity 91.6%; Pred. No. 2.6e-73;
Matches 175; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFPGGGQIVGGVLLPRRGLGVRAIRKTSERSQPRG 60
DB 1 MSTNPKPKQKTKRNTNRRPQDVKFGGGQIVGGVLLPRRGLGVRAIRKTSERSQPRG 60
QY 61 RROPIPKARQPOGRHWAQPGYWPVLYGNEGCGWAGLLSPGRSRPHWGNDRPRRSRLG 120
DB 61 RROPIPKARPEGRWAQPGYWPVLYGNEGCGWAGLLSPGRSRPHWGNDRPRRSRLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2004, 01:34:01 ; Search time 195 Seconds

(without alignments)
563.572 Million cell updates/sec

Title: US-09-084-691B-206

Perfect score: 1045

Sequence: 1 MSTLPKPKQKRNTRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1045	100.0	191	2	Q68115
2	1035	99.0	191	2	Q68411
3	1025	98.1	414	2	P89956
4	1024	98.0	414	2	P89958
5	1024	98.0	414	2	Q81329
6	1024	98.0	3016	2	O92531
7	1023	97.9	3018	2	O39927
8	1020	97.6	414	2	P89963
9	1016	97.2	414	2	P89957
10	1016	97.2	414	2	P89960
11	1014	97.0	414	2	P89959
12	1011	96.7	414	2	Q81270
13	1009	96.6	319	2	Q68703
14	1009	96.6	414	2	P89962
15	1007	96.4	191	2	O93057
16	1007	96.4	414	2	Q81269
17	1003	96.0	414	2	Q81268
18	1003	96.0	414	2	Q81314
19	1001	95.8	191	2	O70635
20	1000	95.7	414	2	Q81271
21	999	95.6	414	2	P89955
22	997	95.4	414	2	P89954
23	997	95.4	415	2	P89961
24	997	95.4	415	2	P89964
25	997	95.4	415	2	P89965
26	997	95.4	3015	2	O92532
27	996	95.3	191	2	Q68153
28	995	95.2	415	2	Q81272
29	995	95.2	415	2	Q81315
30	990	94.7	415	2	P87751
31	990	94.7	3013	2	O92530

```

32 988 94.5 414 2 Q81267 081267 hepatitis c
33 988 94.5 514 2 Q6XJ47 Q6XJ47 hepatitis c
34 988 94.5 514 2 AAB69953 AAB69953 hepatitis c
35 988 94.5 3019 2 O92529 O92529 hepatitis c
36 987 94.4 191 2 P87841 P87841 hepatitis c
37 987 94.4 415 2 Q81274 Q81274 hepatitis c
38 985 94.3 416 2 O39647 O39647 hepatitis c
39 984 94.2 319 2 Q68698 Q68698 hepatitis c
40 984 94.2 415 2 Q81273 Q81273 hepatitis c
41 982 94.0 191 2 Q68152 Q68152 hepatitis c
42 982 94.0 191 2 Q68157 Q68157 hepatitis c
43 982 94.0 319 2 Q68705 Q68705 hepatitis c
44 981 93.9 191 2 Q68116 Q68116 hepatitis c
45 981 93.9 191 2 Q68576 Q68576 hepatitis c

```

ALIGNMENTS

```

RESULT 1
Q68115 PRELIMINARY; PRT; 191 AA.
AC Q68115;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP SEQUENCE FROM N.A.
RA MEDLINE=94336721; PubMed=8058787;
RX Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the core gene of 14 hepatitis C virus
RT genotypes.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:8239-8243(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA Bukh J., Purcell R.H., Miller R.H.;
RT "At least 12 genotypes of hepatitis C virus predicted by sequence
RT analysis of the putative E1 gene of isolates collected worldwide.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:8234-8238(1993).
DR EMBL; U0198; AAA21037.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
DR Polyprotein.
FT CHAIN 1 >191 core protein.
FT NON_TER 191 191
SQ SEQUENCE 191 AA; 20725 MW; 7A7994DF2F0909C6 CRC64;

```

Query Match 100.0%; Score 1045; DB 2; Length 191;

Best Local Similarity 100.0%; Pred. No. 7e-80;

Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKRNTRPTDVKFPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRG 60

DB 1 MSTLPKPKQKRNTRPTDVKFPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRG 60

QY 61 RROPIPKARQPGRHWAQPGVWPPLYGNCGWACWLLSPRGSRPHWGPNDRPRSRNLG 120

```
Db 61 RRQIPKARQPGQRHWAQPGYPWPLYGCGWAGWLLSPRGSRPHWGPNDPRRSRNLG 120
Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 2
Q68411
ID Q68411 PRELIMINARY; PRT; 191 AA.
AC Q68411;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus type 6.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=42182;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=97052554; PubMed=8897188;
RA Bernier L., Willems B., Delage G., Murphy D.G.;
RT "Identification of numerous hepatitis C virus genotypes in Montreal,
Canada.";
RL J. Clin. Microbiol. 34:2815-2818 (1996).
DR EMBL; U33435; AAB40038.1; -.
GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_core; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON TER 191
SQ SEQUENCE 191 AA; 20789 MW; F292AB64B56DE30A CRC64;

Query Match 99.0%; Score 1035; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 4.8e-79;
Matches 189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQRTKNTNRPTDVKPPGGQIVGGVYLLPRGPRLGVRATRTKTSRSQPRG 60
Db 1 MSTLPKPKQRTKNTNRPTDVKPPGGQIVGGVYLLPRGPRLGVRATRTKTSRSQPRG 60
Qy 61 RRQIPKARQPGQRHWAQPGYPWPLYGCGWAGWLLSPRGSRPHWGPNDPRRSRNLG 120
Db 61 RRQIPKARQPGQRHWAQPGYPWPLYGCGWAGWLLSPRGSRPHWGPNDPRRSRNLG 120
Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 3
P89956
ID P89956 PRELIMINARY; PRT; 414 AA.
AC P89956;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;

[2]
```

```
RN SEQUENCE FROM N.A.
RP MEDLINE=95062197; PubMed=7972001;
RX Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
seventh, eighth, and ninth major genetic groups.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Okamoto H.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D88469; BAA13618.1; -.
PIR; P00804; P00804.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Coats protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polypeptide; Transmembrane.
FT NON TER 414
SQ SEQUENCE 414 AA; 44528 MW; 7A3BE1710311C017 CRC64;

Query Match 98.1%; Score 1025; DB 2; Length 414;
Best Local Similarity 97.9%; Pred. No. 7.3e-78;
Matches 187; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQRTKNTNRPTDVKPPGGQIVGGVYLLPRGPRLGVRATRTKTSRSQPRG 60
Db 1 MSTLPKPKQRTKNTNRPTDVKPPGGQIVGGVYLLPRGPRLGVRATRTKTSRSQPRG 60
Qy 61 RRQIPKARQPGQRHWAQPGYPWPLYGCGWAGWLLSPRGSRPHWGPNDPRRSRNLG 120
Db 61 RRQIPKARQPGQRHWAQPGYPWPLYGCGWAGWLLSPRGSRPHWGPNDPRRSRNLG 120
Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 4
P89958
ID P89958 PRELIMINARY; PRT; 414 AA.
AC P89958;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95062197; PubMed=7972001;
RX Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
seventh, eighth, and ninth major genetic groups.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026 (1994).
RN [2]
```

```

RP SEQUENCE FROM N.A.
RA Okamoto H.; Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; D88471; BAA13620.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NSI.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NSI; 1.
DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 414
SQ SEQUENCE 414 AA; 44973 MW; F3F3CF154372F4FF CRC64;

Query Match 98.0%; Score 1024; DB 2; Length 414;
Best Local Similarity 97.4%; Pred. No. 8.8e-78;
Matches 186; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFPGGQIVGGVLLPRGRLGVRATKTSERSQPRG 60
DB 1 MSTLPKPKQKTKRNTNRRPMDVKFPGGQIVGGVLLPRGRLGVRATKTSERSQPRG 60

QY 61 RRQIPKARQPOGRHWAOPGYWPPLYGNEGCGWAGLLSPGRSRPHWGPNDRRRSRNLG 120
DB 61 RRQIPKARQPTGRHWAOPGYWPPLYGNEGCGWAGLLSPGRSRPHWGPNDRRRSRNLG 120

QY 121 KVIDLTTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVIDLTTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 5
Q81329 PRELIMINARY; PRT; 414 AA.
AC Q81329;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda P., Nakata S., Chosa T.,
RA Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
RT seventh, eighth, and ninth major genetic groups.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Okamoto H.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D88469; BAA13617.1; -.
DR PIR; PQ0804; PQ0804.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFPGGQIVGGVLLPRGRLGVRATKTSERSQPRG 60
DB 1 MSTLPKPKQKTKRNTNRRPMDVKFPGGQIVGGVLLPRGRLGVRATKTSERSQPRG 60

QY 61 RRQIPKARQPOGRHWAOPGYWPPLYGNEGCGWAGLLSPGRSRPHWGPNDRRRSRNLG 120
DB 61 RRQIPKARQPTGRHWAOPGYWPPLYGNEGCGWAGLLSPGRSRPHWGPNDRRRSRNLG 120

QY 121 KVIDLTTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVIDLTTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 6
Q92531 PRELIMINARY; PRT; 3016 AA.
AC Q92531;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98378034; PubMed=9714232;
RA Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda P., Miyakawa Y.,
RA Mayumi M.;
RT "The entire nucleotide sequences of three hepatitis C virus isolates
RT in genetic groups 7-9 and comparison with those in the other eight
RT genetic groups.";
RL J. Gen. Virol. 79:1847-1857(1998).
DR EMBL; D84264; BAA32666.1; -.
DR PIR; PQ0804; PQ0804.
DR HSP; P27958; 1AIV.
DR MEROPS; S29.001; -.
DR MEROPS; U39.001; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0008508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.

```

DR InterPro: IPR0011410; DEAD.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRP.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR009003; Pept Ser Cys.
DR InterPro: IPR002518; Pept U39 HCV NS2.
DR InterPro: IPR007095; RNA pol DS PS.
DR InterPro: IPR007094; RNA pol PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01538; HCV NS1; 1.
DR Pfam: PF01537; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00998; Viral RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3016 AA; 328032 MW; 4ESCFF96258BCE3B CRC64;

Query Match 98.0%; Score 1024; DB 2; Length 3016;
Best Local Similarity 98.4%; Pred. No. 6.6e-77;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVRAATRTKTSRSQPRG 60
Db 1 MSTLPKPKQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVRAATRTKTSRSQPRG 60

Qy 61 RRQIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGNPDPRRSRNIG 120
Db 61 RRQIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGNPDPRRSRNIG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

Query Match 97.9%; Score 1023; DB 2; Length 3018;
Best Local Similarity 97.9%; Pred. No. 8e-77;
Matches 187; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVRAATRTKTSRSQPRG 60
Db 1 MSTLPKPKQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVRAATRTKTSRSQPRG 60

Qy 61 RRQIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGNPDPRRSRNIG 120
Db 61 RRQIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGNPDPRRSRNIG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 8
P89963 PRELIMINARY; PRT; 414 AA.
ID P89963
AC P89963
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RA "Hepatitis C virus variants from Vietnam are classifiable into the
RT seventh, eighth, and ninth major genetic groups.";
RT Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
RL [2]
RN SEQUENCE FROM N.A.
RP Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RX EMBL; D88470; BAA13625.1; -;
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0019028; C: viral capsid; IEA.
DR GO: 0019031; C: viral envelope; IEA.
DR GO: 0005198; F: structural molecule activity; IEA.
DR Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Okamoto H.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RX EMBL; D88470; BAA13625.1; -;
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0019028; C: viral capsid; IEA.
DR GO: 0019031; C: viral envelope; IEA.
DR GO: 0005198; F: structural molecule activity; IEA.
DR InterPro: IPR009100; AcylCoA dehyd_NM.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV env.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Coar protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 414
SQ SEQUENCE 414 AA; 44804 MW; 397D0DC5C4BDA20D CRC64;

Query Match 97.6%; Score 1020; DB 2; Length 414;
Best Local Similarity 96.9%; Pred. No. 1.9e-77;
Matches 185; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVLLPRRGLGVRATKTSERSQPRG 60
DB 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVLLPRRGLGVRATKTSERSQPRG 60

QY 61 RRQIPKARQPGQRHWAQPGYPMPLVYGNCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
DB 61 RRQIPKARQPGQRHWAQPGYPMPLVYGNCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120

QY 121 KVITDITCGFADLMGYPVVGAPLGGVAAALAHGVAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVITDITCGFADLMGYPVVGAPLGGVAAALAHGVAIEDGINYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 9
P89957
ID P89957 PRELIMINARY; PRT; 414 AA.
AC P89957;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;

DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;

RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
seventh, eighth, and ninth major genetic groups.";
Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Okamoto H.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RX EMBL; D88470; BAA13619.1; -;
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0019028; C: viral capsid; IEA.
DR GO: 0019031; C: viral envelope; IEA.
DR GO: 0005198; F: structural molecule activity; IEA.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Coar protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 414
SQ SEQUENCE 414 AA; 44850 MW; 1754DD583F52414 CRC64;

Query Match 97.2%; Score 1016; DB 2; Length 414;
Best Local Similarity 96.9%; Pred. No. 4.1e-77;
Matches 185; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVLLPRRGLGVRATKTSERSQPRG 60
DB 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVLLPRRGLGVRATKTSERSQPRG 60

QY 61 RRQIPKARQPGQRHWAQPGYPMPLVYGNCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
DB 61 RRQIPKARQPGQRHWAQPGYPMPLVYGNCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120

QY 121 KVITDITCGFADLMGYPVVGAPLGGVAAALAHGVAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVITDITCGFADLMGYPVVGAPLGGVAAALAHGVAIEDGINYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 10
P89960
ID P89960 PRELIMINARY; PRT; 414 AA.
AC P89960;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
seventh, eighth, and ninth major genetic groups.";
Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Okamoto H.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RX EMBL; D88473; BAA13622.1; -;

```

DR PIR; PQ0804; PQ0804.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01539; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
FT NON TER 414
SQ SEQUENCE 414 AA; 44630 MW; C7DDC84D01988BD7 CRC64;

Query Match 97.2%; Score 1016; DB 2; Length 414;
Best Local Similarity 96.9%; Pred. No. 4.1e-77;
Matches 185; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MSTLPKQPKTKRNTNRPTDVKPPGGQIVGGYLLPRRGPRLGVRATKTSERSQPRG 60
Db 1 MSTLPKQPKTKRNTNRPTDVKPPGGQIVGGYLLPRRGPRLGVRATKTSERSQPRG 60
Qy 61 RRQPIKARQPGRRHWAQPGYPWPLYGNEGCGWAGLLSPRGRPHGPNDRPRRSNIG 120
Db 61 RRQPIKARQPGRRHWAQPGYPWPLYGNEGCGWAGLLSPRGRPHGPNDRPRRSNIG 120
Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 11
ID P89959 PRELIMINARY; PRT; 414 AA.
AC P89959;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
seventh, eighth, and ninth major genetic groups.";
Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
RL [2]
RN SEQUENCE FROM N.A.
RA Okamoto H.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
EMBL; D88472; BAA13621.1; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01543; HCV capsid; 1.

Qy 1 MSTLPKQPKTKRNTNRPTDVKPPGGQIVGGYLLPRRGPRLGVRATKTSERSQPRG 60
Db 1 MSTLPKQPKTKRNTNRPTDVKPPGGQIVGGYLLPRRGPRLGVRATKTSERSQPRG 60
Qy 61 RRQPIKARQPGRRHWAQPGYPWPLYGNEGCGWAGLLSPRGRPHGPNDRPRRSNIG 120
Db 61 RRQPIKARQPGRRHWAQPGYPWPLYGNEGCGWAGLLSPRGRPHGPNDRPRRSNIG 120
Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 12
ID Q81270 PRELIMINARY; PRT; 414 AA.
AC Q81270;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96005057; PubMed=7561773;
RA Tokita H., Okamoto H., Luengrojanakul P., Varesangthip K.,
Chainuvati T., Iizuka H., Tsuda F., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Thailand classifiable into five novel
genotypes in the sixth (6b), seventh (7c,7d) and ninth (9b, 9c) major
genetic groups.";
J. Gen. Virol. 76:2329-2335(1995).
RL EMBL; D37845; BAA07091.1; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
FT NON TER 414
SQ SEQUENCE 414 AA; 44994 MW; 8C1E123FC61D89FC CRC64;

Query Match 96.7%; Score 1011; DB 2; Length 414;
Best Local Similarity 97.4%; Pred. No. 1.1e-76;
Matches 186; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MSTLPKQPKTKRNTNRPTDVKPPGGQIVGGYLLPRRGPRLGVRATKTSERSQPRG 60

```

```

Db      1 MSTLPKQKTKNTNRRPTDVKFGGQIVGGVILLPRGRLGVRATKTSERSQPRG 60
QY      61 RRQPIPKARQPOGRHWAQPGYWPVLYGNECGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
Db      61 RRQPIPKARHQTGRSQAQPGYWPVLYGNECGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
QY      121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db      121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY      181 LLSCLTTPASA 191
Db      181 LLSCLTTPASA 191

RESULT 13
Q68703
ID      068703      PRELIMINARY;      PRT;      319 AA.
AC      068703;
DT      01-NOV-1996 (TReMBLrel. 01, Created)
DT      01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE      01-JUN-2003 (TReMBLrel. 24, Last annotation update)
OS      Core protein (Fragment).
OC      Hepatitis C virus type 6a.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Hepacivirus; Hepatitis C virus type 6.
OX      NCBI_TaxID=31655;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96118171; PubMed=8578855;
RA      Stuyver L., Wyseur A., van Arnhem W., Lunel F., Laurent-Puig P.,
RA      Pawlotsky J.-M., Klefer B., Basit L., Nkengasong J., van Doorn L.-J.,
RA      Maertens G.;
RT      "Hepatitis C virus genotyping by means of 5'-UR/core line probe assays
RT      and molecular analysis of untypeable samples.";
RL      Virus Res. 38:137-157(1995).
DR      EMBL; L38339; AAC42193.1; -
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0019031; C:viral envelope; IEA.
DR      GO; GO:0005198; F:structural molecule activity; IEA.
DR      InterPro; IPR002522; HCV capsid.
DR      InterPro; IPR002521; HCV core.
DR      Pfam; PF01543; HCV_capsid; 1.
DR      Pfam; PF01539; HCV_core; 1.
DR      Pfam; PF01539; HCV env; 1.
KW      Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW      Transmembrane.
FT      CHAIN      1      191      core protein.
FT      NON_TER    192      319      E1 protein.
SQ      SEQUENCE 319 AA; 34545 MW; DC1D0EA52ED64A8F CRC64;

Query Match      96.6%; Score 1009; DB 2; Length 319;
Best Local Similarity 96.3%; Pred. No. 1.2e-76;
Matches 184; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      1 MSTLPKQKTKNTNRRPTDVKFGGQIVGGVILLPRGRLGVRATKTSERSQPRG 60
Db      1 MSTLPKQKTKNTNRRPTDVKFGGQIVGGVILLPRGRLGVRATKTSERSQPRG 60
QY      61 RRQPIPKARQPOGRHWAQPGYWPVLYGNECGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
Db      61 RRQPIPKARQTOGRHWAQPGYWPVLYGNECGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
QY      121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db      121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNITGCSFSIFLLA 180
QY      181 LLSCLTTPASA 191
Db      181 LLSCLTTPASA 191

```

```

Db      181 LLSCLTTPASA 191

RESULT 14
P89962
ID      P89962      PRELIMINARY;      PRT;      414 AA.
AC      P89962;
DT      01-MAY-1997 (TReMBLrel. 03, Created)
DT      01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT      01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE      Core, env and part of E2/NS1 (Fragment).
OS      Hepatitis C virus.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Hepacivirus.
OX      NCBI_TaxID=11103;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=95062197; PubMed=7972001;
RA      Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA      Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.;
RT      "Hepatitis C virus variants from Vietnam are classifiable into the
RT      seventh, eighth, and ninth major genetic groups.";
RL      Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Okamoto H.;
RL      Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR      EMBL; D88475; BAAL3624.1; -
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0019031; C:viral envelope; IEA.
DR      GO; GO:0005198; F:structural molecule activity; IEA.
DR      InterPro; IPR002522; HCV_capsid.
DR      InterPro; IPR002521; HCV core.
DR      InterPro; IPR002519; HCV env.
DR      InterPro; IPR002531; HCV NS1.
DR      Pfam; PF01543; HCV_capsid; 1.
DR      Pfam; PF01542; HCV_core; 1.
DR      Pfam; PF01539; HCV env; 1.
DR      Pfam; PF01560; HCV NS1; 1.
KW      Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW      Polyprotein; Transmembrane.
FT      NON_TER    414      414
SQ      SEQUENCE 414 AA; 44751 MW; 4BBA61438D11E3C6 CRC64;

Query Match      96.6%; Score 1009; DB 2; Length 414;
Best Local Similarity 96.3%; Pred. No. 1.6e-76;
Matches 184; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      1 MSTLPKQKTKNTNRRPTDVKFGGQIVGGVILLPRGRLGVRATKTSERSQPRG 60
Db      1 MSTLPKQKTKNTNRRPTDVKFGGQIVGGVILLPRGRLGVRATKTSERSQPRG 60
QY      61 RRQPIPKARQPOGRHWAQPGYWPVLYGNECGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
Db      61 RRQPIPKARQSQGRHWAQPGYWPVLYGNECGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
QY      121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db      121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINFATGNLPGCSFSIFLLA 180
QY      181 LLSCLTTPASA 191
Db      181 LLSCLTTPASA 191

RESULT 15
O93057
ID      O93057      PRELIMINARY;      PRT;      191 AA.
AC      O93057; Q68747;
DT      01-NOV-1998 (TReMBLrel. 08, Created)
DT      01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT      01-JUN-2003 (TReMBLrel. 24, Last annotation update)

```

```
DE Polyprotein (Fragment).
OC Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97060095; PubMed=8904416;
RA Doi H.; Apichartpiyakul C.; Ohba K.; Mizokami M.; Hotta H.;
RT "Hepatitis C virus (HCV) subtype prevalence in Chiang Mai, Thailand,
RL and identification of novel subtypes of HCV major types 6."
J. Clin. Microbiol. 34:569-574(1996).
DR EMBL; D63946; BAA09976.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein.
FT CHAIN 1 >191 core protein.
FT NON_TER 191 191
FT SEQUENCE 191 AA; 20727 MW; 411C1C851ADA2DB5 CRC64;

Query Match 96.4%; Score 1007; DB 2; Length 191;
Best Local Similarity 95.3%; Pred. No. 1.1e-76;
Matches 182; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQKTKRNTNRPTDKFPFGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Db |||||
1 MSTLPKPKQKTKRNTNRPMDDVKFPFGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Qy 61 RRQPIKARQPGRHWAQPGYPWPLYGNEGCGWAGLLSPRGSRPHMGPNDRPRSRNLG 120
Db |||||
61 RRQPIKARQPTGRHWAQPGYPWPLYGNEGCGWAGLLSPRGSRPNWGPNDPRRRNLG 120
Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db :|||:
121 RVIDTITCGLADLMGYIPVLGAPLGVAALAHGVVVEDGINYATGNLPGCSFSIFLLA 180
Qy 181 LLSCLITTPASA 191
Db |||||
181 LLSCLITTPASA 191

Search completed: October 30, 2004, 01:46:04
Job time : 197 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2004, 01:36:27 ; Search time 39 Seconds
(without alignments)
324.788 Million cell updates/sec

Title: US-09-084-691B-206

Perfect score: 1045

Sequence: 1 MSTLPKPKYKNTNRPT.....CSFSFLALLSCLTPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/aaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/aaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/aaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/aaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/aaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1045	100.0	191	2	US-08-290-665A-206
2	1045	100.0	191	5	PCT-US95-10398-206
3	1045	100.0	319	4	US-08-635-886C-228
4	1045	100.0	319	4	US-08-974-690C-228
5	996	95.3	191	2	US-08-290-665A-191
6	996	95.3	191	5	PCT-US95-10398-191
7	982	94.0	191	2	US-08-290-665A-192
8	982	94.0	191	2	US-08-290-665A-193
9	982	94.0	191	5	PCT-US95-10398-192
10	982	94.0	191	5	PCT-US95-10398-193
11	981	93.9	191	2	US-08-290-665A-172
12	981	93.9	191	5	PCT-US95-10398-172
13	980	93.8	191	2	US-08-290-665A-197
14	980	93.8	191	5	PCT-US95-10398-197
15	980	93.8	319	4	US-08-635-886C-217
16	980	93.8	319	4	US-08-974-690C-217
17	978	93.6	191	2	US-08-290-665A-195
18	978	93.6	191	5	PCT-US95-10398-195
19	977	93.5	191	2	US-08-290-665A-196
20	977	93.5	191	5	PCT-US95-10398-196
21	977	93.5	319	3	US-08-836-075A-42
22	977	93.5	319	4	US-08-635-886C-229
23	977	93.5	319	4	US-08-974-690C-229
24	977	93.5	450	4	US-08-635-886C-181
25	977	93.5	450	4	US-08-974-690C-181
26	977	93.5	2894	2	US-08-466-975A-23
27	977	93.5	2894	2	US-08-391-671A-23

28 977 93.5 2894 3 US-08-467-902A-23 Sequence 23, Appl
29 977 93.5 2894 3 US-09-275-265-23 Sequence 23, Appl
30 977 93.5 2894 4 US-09-941-611-23 Sequence 23, Appl
31 976 93.4 191 2 US-08-290-665A-156 Sequence 156, Appl
32 976 93.4 191 2 US-08-290-665A-157 Sequence 157, Appl
33 976 93.4 191 2 US-08-290-665A-158 Sequence 158, Appl
34 976 93.4 191 2 US-08-290-665A-159 Sequence 159, Appl
35 976 93.4 191 2 US-08-290-665A-160 Sequence 160, Appl
36 976 93.4 191 3 US-08-380-160-3 Sequence 3, Appl
37 976 93.4 191 5 PCT-US95-10398-156 Sequence 156, Appl
38 976 93.4 191 5 PCT-US95-10398-157 Sequence 157, Appl
39 976 93.4 191 5 PCT-US95-10398-158 Sequence 158, Appl
40 976 93.4 191 5 PCT-US95-10398-159 Sequence 159, Appl
41 976 93.4 191 5 PCT-US95-10398-160 Sequence 160, Appl
42 976 93.4 450 4 US-08-635-886C-179 Sequence 179, Appl
43 976 93.4 450 4 US-08-635-886C-180 Sequence 180, Appl
44 976 93.4 450 4 US-08-974-690C-179 Sequence 179, Appl
45 976 93.4 450 4 US-08-974-690C-180 Sequence 180, Appl

ALIGNMENTS

RESULT 1
US-08-290-665A-206
; Sequence 206, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 751-6800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 206:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: HK2
; US-08-290-665A-206

Query Match 100.0%; Score 1045; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 9.3e-97;

Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVTRATKTSRSOPRG 60
Db 1 MSTLPKPKQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVTRATKTSRSOPRG 60

Qy 61 RRQIPKARQPGQRHWAQPGYPWPPLYNCGCGWAGWLLSPRSGRPHGPNDRRRSRNLG 120
Db 61 RRQIPKARQPGQRHWAQPGYPWPPLYNCGCGWAGWLLSPRSGRPHGPNDRRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 2

PCT-US95-10398-206
; Sequence 206, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BURK, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 206:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homocapiens
; INDIVIDUAL ISOLATE: HK2
PCT-US95-10398-206

Query Match

100.0%; Score 1045; DB 5; Length 191;

Best Local Similarity 100.0%; Pred. No. 9.3e-97;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVTRATKTSRSOPRG 60
Db 1 MSTLPKPKQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVTRATKTSRSOPRG 60

Qy 61 RRQIPKARQPGQRHWAQPGYPWPPLYNCGCGWAGWLLSPRSGRPHGPNDRRRSRNLG 120
Db 61 RRQIPKARQPGQRHWAQPGYPWPPLYNCGCGWAGWLLSPRSGRPHGPNDRRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 3

US-08-635-886C-228
; Sequence 228, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 228
; LENGTH: 319
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-228

Query Match

100.0%; Score 1045; DB 4; Length 319;

Best Local Similarity 100.0%; Pred. No. 1.7e-96;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVTRATKTSRSOPRG 60
Db 1 MSTLPKPKQRTKNTNRPTDVKFPGGGQIVGGVYLLPRGPRGLGVTRATKTSRSOPRG 60

Qy 61 RRQIPKARQPGQRHWAQPGYPWPPLYNCGCGWAGWLLSPRSGRPHGPNDRRRSRNLG 120
Db 61 RRQIPKARQPGQRHWAQPGYPWPPLYNCGCGWAGWLLSPRSGRPHGPNDRRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 4

US-08-974-690C-228
; Sequence 228, Application US/08974690C
; Patent No. 661333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert

APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
FILE REFERENCE: 2551-94
CURRENT APPLICATION NUMBER: US/08/974,690C
CURRENT FILING DATE: 1997-11-19
PRIOR APPLICATION NUMBER: PCT/EP94/03555
PRIOR FILING DATE: 1994-10-28
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PatentIn version 3.1
SEQ ID NO 228
LENGTH: 319
TYPE: PRT
ORGANISM: hepatitis C virus
US-08-974-690C-228

Query Match 100.0%; Score 1045; DB 4; Length 319;
Best Local Similarity 100.0%; Pred. No. 1.7e-96;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSTLPKPKORTKRNTRPTDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
DB 1 MSTLPKPKORTKRNTRPTDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
QY 61 RRQPIPKARQPGRHWAQPGYPWPLYGNEGCWAGWLLSPRGRPHWGPNDRPRRSNIG 120
DB 61 RRQPIPKARQPGRHWAQPGYPWPLYGNEGCWAGWLLSPRGRPHWGPNDRPRRSNIG 120
QY 121 KVIDTLTCGFADLMGYIPVWGAPLGCVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVWGAPLGCVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 5

US-08-290-665A-191
Sequence 191, Application US/08290665A
Patent No. 582852
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 191:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: Z4
US-08-290-665A-191

Query Match 95.3%; Score 996; DB 2; Length 191;
Best Local Similarity 94.8%; Pred. No. 7.3e-92;
Matches 181; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 1 MSTLPKPKORTKRNTRPTDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
DB 1 MSTLPKPKORTKRNTRPTDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
QY 61 RRQPIPKARQPGRHWAQPGYPWPLYGNEGCWAGWLLSPRGRPHWGPNDRPRRSNIG 120
DB 61 RRQPIPKARQPGRHWAQPGYPWPLYGNEGCWAGWLLSPRGRPHWGPNDRPRRSNIG 120
QY 121 KVIDTLTCGFADLMGYIPVWGAPLGCVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVWGAPLGCVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 6

PCT-US95-10398-191
Sequence 191, Application PC/TUS9510398
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 191:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: Z4
PCT-US95-10398-191

Query Match 95.3%; Score 996; DB 5; Length 191;
Best Local Similarity 94.8%; Pred. No. 7.3e-92;
Matches 181; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MSTLPKQRTKNTNRPTDVKFPGGQIVGGVYLLPRRGLRVATRTKTSRSQPRG 60
Db 1 MSTNPKQRTKNTNRPMQVDFPGGQIVGGVYLLPRRGLRVATRTKTSRSQPRG 60

Qy 61 RRQIPKARQPGHWAQPGYPWPLYNCGCGWAGWLLSPRGSRPHGPNDRPRRSNGLG 120
Db 61 RRQIPKARQPGHWAQPGYPWPLYNCGCGWAGWLLSPRGSRPHGPNDRPRRSNGLG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRATIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRATIEDGINVATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 7
US-08-290-665A-192
Sequence 192, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792

INFORMATION FOR SEQ ID NO: 192:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: Z8
US-08-290-665A-192

Query Match 94.0%; Score 982; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 1.8e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MSTLPKQRTKNTNRPTDVKFPGGQIVGGVYLLPRRGLRVATRTKTSRSQPRG 60
Db 1 MSTNPKQRTKNTNRPMQVDFPGGQIVGGVYLLPRRGLRVATRTKTSRSQPRG 60

Qy 61 RRQIPKARQPGHWAQPGYPWPLYNCGCGWAGWLLSPRGSRPHGPNDRPRRSNGLG 120
Db 61 RRQIPKARQPGHWAQPGYPWPLYNCGCGWAGWLLSPRGSRPHGPNDRPRRSNGLG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRATIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRATIEDGINVATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 8
US-08-290-665A-193
Sequence 193, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid

STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: Z1
US-08-290-665A-193

Query Match 94.0%; Score 982; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 1.8e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRPTDVKFPGGGQIVGGVYLLPRRGRLGVRAATKTSERSQPRG 60
DB 1 MSTNPKPKQKTKRNTNRPTDVKFPGGGQIVGGVYLLPRRGRLGVRAATKTSERSQPRG 60
QY 61 RRQPIKARQPGRIWAQPGYPWPPLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSNIG 120
DB 61 RRQPIKARRSSEGRSWAQPGYPWPPLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSNIG 120
QY 121 KVIDLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
DB 121 KVIDLTCGFADLMGYIPLVGAPVGGVARALAHGVRAVEDGINVATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 9

PCT-US95-10398-192
Sequence 192, Application PC/TUS9510398
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R. H. AND
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSER: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 192:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids

TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: Z8
PCT-US95-10398-192

Query Match 94.0%; Score 982; DB 5; Length 191;
Best Local Similarity 93.7%; Pred. No. 1.8e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRPTDVKFPGGGQIVGGVYLLPRRGRLGVRAATKTSERSQPRG 60
DB 1 MSTNPKPKQKTKRNTNRPTDVKFPGGGQIVGGVYLLPRRGRLGVRAATKTSERSQPRG 60
QY 61 RRQPIKARQPGRIWAQPGYPWPPLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSNIG 120
DB 61 RRQPIKARRSSEGRSWAQPGYPWPPLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSNIG 120
QY 121 KVIDLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
DB 121 KVIDLTCGFADLMGYIPLVGAPVGGVARALAHGVRAVEDGINVATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 10

PCT-US95-10398-193
Sequence 193, Application PC/TUS9510398
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R. H. AND
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSER: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:

LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
ORGANISM: hominid
INDIVIDUAL ISOLATE: Z1
PCT-US95-10398-193

Query Match 94.0%; Score 982; DB 5; Length 191;
Best Local Similarity 93.7%; Pred. No. 1.8e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 1 MSTLPKQKTKRNTNRRPTDVKPPGGQIVGGVYLLPRRGPRLGVRAATKTSRSQPRG 60
DB 1 MSTNPKQKTKRNTNRRPTDVKPPGGQIVGGVYLLPRRGPRLGVRAATKTSRSQPRG 60
QY 61 RRQIPKARQOQGRHWAQPGYPWLYNGEGCGWAGWLLSPRGSPPHMGPNDRSRNLG 120
DB 61 RRQIPKARQOQGRHWAQPGYPWLYNGEGCGWAGWLLSPRGSPPHMGPNDRSRNLG 120
QY 121 KVIDLTGCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
DB 121 KVIDLTGCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 11
US-08-290-665A-172
Sequence 172, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
APPLICANT: BURKH, J., MILLER, R.H. AND
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 172:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown

ORIGINAL SOURCE:
ORGANISM: hominid
INDIVIDUAL ISOLATE: HK3
US-08-290-665A-172

Query Match 93.9%; Score 981; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 2.3e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 1 MSTLPKQKTKRNTNRRPTDVKPPGGQIVGGVYLLPRRGPRLGVRAATKTSRSQPRG 60
DB 1 MSTNPKQKTKRNTNRRPTDVKPPGGQIVGGVYLLPRRGPRLGVRAATKTSRSQPRG 60
QY 61 RRQIPKARQOQGRHWAQPGYPWLYNGEGCGWAGWLLSPRGSPPHMGPNDRSRNLG 120
DB 61 RRQIPKARQOQGRHWAQPGYPWLYNGEGCGWAGWLLSPRGSPPHMGPNDRSRNLG 120
QY 121 KVIDLTGCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
DB 121 KVIDLTGCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 12
PCT-US95-10398-172
Sequence 172, Application PC/TUS9510398
GENERAL INFORMATION:
APPLICANT: BURKH, J., MILLER, R.H. AND
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 172:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown

```
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: HK3
PCT-US95-10398-172

Query Match      93.9%; Score 981; DB 5; Length 191;
Best Local Similarity 93.7%; Pred. No. 2.3e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSTLPKPKRKTNTNRRPTDVKFFGGQIVGGVILLPRGRPLGVRAATKTSERSQPRG 60
   |||||
Db 1 MSTNPKPKRKTNTNRRPTDVKFFGGQIVGGVILLPRGRPLGVRAATKTSERSQPRG 60
   |||||

QY 61 RROPIPKARQPOGRHWAQPGYPMPLYGNEGCWAGWLLSPRGRPHWGNDPRRNRNLG 120
   |||||
Db 61 RROPIPKARQPOGRHWAQPGYPMPLYGNEGCWAGWLLSPRGRPHWGNDPRRNRNLG 120
   |||||

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
   |||||
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
   |||||

QY 181 LLSCLTTPASA 191
   |||||
Db 181 LLSCLTTPASA 191
   |||||

RESULT 13
US-08-290-665A-197
; Sequence 197, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 197:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: DK13

US-08-290-665A-197
Query Match      93.8%; Score 980; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 2.9e-90;
Matches 179; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSTLPKPKRKTNTNRRPTDVKFFGGQIVGGVILLPRGRPLGVRAATKTSERSQPRG 60
   |||||
Db 1 MSTNPKPKRKTNTNRRPTDVKFFGGQIVGGVILLPRGRPLGVRAATKTSERSQPRG 60
   |||||

QY 61 RROPIPKARQPOGRHWAQPGYPMPLYGNEGCWAGWLLSPRGRPHWGNDPRRNRNLG 120
   |||||
Db 61 RROPIPKARQPOGRHWAQPGYPMPLYGNEGCWAGWLLSPRGRPHWGNDPRRNRNLG 120
   |||||

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
   |||||
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
   |||||

QY 181 LLSCLTTPASA 191
   |||||
Db 181 LLSCLTTPASA 191
   |||||

RESULT 14
PCT-US95-10398-197
; Sequence 197, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 197:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2004, 01:40:17 ; Search time 128 Seconds
(without alignments)
483.791 Million cell updates/sec

Title: US-09-084-691B-206
Perfect score: 1045
Sequence: 1 MSTLPKQKTKYKRNRRPT.....CSFSIFILLALLSLCTTPASA 191

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1045	100.0	319	15	US-10-651-165-228
2	980	93.8	319	15	US-10-651-165-228
3	977	93.5	319	9	US-09-851-138-42
4	977	93.5	319	15	US-10-651-165-229
5	977	93.5	450	15	US-10-651-165-181
6	977	93.5	2894	9	US-09-941-611-23
7	977	93.5	2894	14	US-10-044-995-23
8	976	93.4	249	15	US-10-365-620-54
9	976	93.4	450	15	US-10-651-165-179
10	976	93.4	450	15	US-10-651-165-180
11	976	93.4	473	15	US-10-365-620-56
12	976	93.4	3011	9	US-09-742-659-4
13	976	93.4	3011	9	US-09-952-572-9

14	976	93.4	3011	9	US-09-929-955-1	Sequence 1, Appli
15	976	93.4	3011	9	US-09-747-419-20	Sequence 20, Appli
16	976	93.4	3011	10	US-09-891-894-3	Sequence 3, Appli
17	976	93.4	3011	13	US-10-104-366-1	Sequence 1, Appli
18	976	93.4	3011	14	US-10-259-275-20	Sequence 20, Appli
19	976	93.4	3011	14	US-10-184-150-3	Sequence 3, Appli
20	976	93.4	3011	14	US-10-328-997-3	Sequence 14, Appli
21	976	93.4	3011	15	US-10-189-359-14	Sequence 14, Appli
22	976	93.4	3011	15	US-10-296-734-406	Sequence 406, App
23	976	93.4	3011	15	US-10-719-619-1	Sequence 1, Appli
24	976	93.4	3012	9	US-09-238-076-2	Sequence 2, Appli
25	976	93.4	3012	10	US-09-995-937-2	Sequence 2, Appli
26	976	93.4	3012	10	US-09-917-563-2	Sequence 20, Appli
27	972	93.0	3011	9	US-09-238-076-20	Sequence 20, Appli
28	972	93.0	3011	10	US-09-995-937-20	Sequence 20, Appli
29	972	93.0	3011	10	US-09-917-563-20	Sequence 190, App
30	970	92.8	450	15	US-10-651-165-190	Sequence 189, App
31	969	92.7	450	15	US-10-651-165-189	Sequence 1, Appli
32	969	92.7	3010	15	US-10-467-000-1	Sequence 219, App
33	968	92.6	319	15	US-10-651-165-219	Sequence 2, Appli
34	968	92.6	3011	9	US-09-916-359-2	Sequence 3, Appli
35	968	92.6	3011	16	US-10-445-724-2	Sequence 3, Appli
36	967	92.5	191	10	US-09-194-949-3	Sequence 187, App
37	965	92.3	450	15	US-10-651-165-187	Sequence 193, App
38	965	92.3	450	15	US-10-651-165-193	Sequence 1, Appli
39	964	92.2	190	14	US-10-268-562-1	Sequence 48, Appli
40	964	92.2	319	9	US-09-851-138-48	Sequence 231, App
41	964	92.2	319	15	US-10-651-165-231	Sequence 7, Appli
42	963	92.2	190	15	US-10-450-649-7	Sequence 50, Appli
43	962	92.1	809	9	US-09-973-025-50	Sequence 50, Appli
44	962	92.1	809	10	US-09-899-303-50	Sequence 50, Appli
45	962	92.1	809	10	US-09-995-808-50	Sequence 50, Appli

ALIGNMENTS

RESULT 1

US-10-651-165-228
; Sequence 228, Application US/10651165
; Publication No. US2004004787A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 228
; LENGTH: 319
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-228

Query Match 100.0%; Score 1045; DB 15; Length 319;
Best Local Similarity 100.0%; Pred. No. 1.1e-83;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSTLPKQKTKYKRNRRPTDVKFGGGQIVGGVYLLPRGVLGVRAETKTSERSQPRG	60
Db	1	MSTLPKQKTKYKRNRRPTDVKFGGGQIVGGVYLLPRGVLGVRAETKTSERSQPRG	60
Qy	61	RRQPIKARQPGQRHWAQPGYPWPPLYGNECGWAGLLSPRGRPHWPNDRRRRNLG	120

Db 61 RROPIKARQPGRHWAQPGYPWPLYGNEGCWAGWLLSPGRSPHGWGPNDRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGYNATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGYNATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 2

US-10-651-165-217
; Sequence 217, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 217
; LENGTH: 319
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-217

Query Match 93.8%; Score 980; DB 15; Length 319;
Best Local Similarity 93.7%; Pred. No. 5.9e-78;
Matches 179; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 1 MSTLPKPKQRTKRNTRRPDVKFPGGQIVGGVYLLPRGPRGLGVRAATKTSRSQPRG 60
Db 1 MSTNPKPKQRTKRNTRRPDVKFPGGQIVGGVYLLPRGPRGLGVRAATKTSRSQPRG 60
QY 61 RROPIKARQPGRHWAQPGYPWPLYGNEGCWAGWLLSPGRSPHGWGPNDRSRNLG 120
Db 61 RROPIKARQLEGRSNAQPGYPWPLYGNEGCWAGWLLSPGRSPHGWGPNDRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGYNATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGYNATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 3

US-09-851-138-42
; Sequence 42, Application US/09851138
; Publication No. US20020183508A1
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
; AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
; AGENTS
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE

; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word 6.0 / ASCII text output
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,138
; FILING DATE: 09-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,075
; FILING DATE: <Unknown>
; APPLICATION NUMBER: EP 94870166.9
; FILING DATE: 21 Oct 1994
; APPLICATION NUMBER: EP 95870076.7
; FILING DATE: 28 Jun 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KAMMERER, PATRICIA A.
; REGISTRATION NUMBER: 29,775
; REFERENCE/DOCKET NUMBER: INNS:004
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-09-851-138-42

Query Match 93.5%; Score 977; DB 9; Length 319;
Best Local Similarity 93.7%; Pred. No. 1.1e-77;
Matches 179; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 1 MSTLPKPKQRTKRNTRRPDVKFPGGQIVGGVYLLPRGPRGLGVRAATKTSRSQPRG 60
Db 1 MSTLPKPKQRTKRNTRRPDVKFPGGQIVGGVYLLPRGPRGLGVRAATKTSRSQPRG 60
QY 61 RROPIKARQPGRHWAQPGYPWPLYGNEGCWAGWLLSPGRSPHGWGPNDRSRNLG 120
Db 61 RROPIKARQPTGRSWGQPGYPWPLYGNEGCWAGWLLSPGRSPHGWGPNDRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGYNATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTXLADLMGYIPVLGFLGGVAAALAHGVRAIEDGVNATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 4

US-10-651-165-229
; Sequence 229, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04

RESULT 6

RESULT 6
US-09-941-611-23
; Sequence 23, Application US/09941611
; Patent No. US20020106640A1

GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
POLLET, DIRK
MAERTENS, GEERT
VAN HEUVERSWUN, HUGO
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
ANTIBODIES TO HEPATITIS C VIRUS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA

APPLICATION NUMBER: US/09/941,611
FILING DATE: 30-AUG-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/391,671
FILING DATE: 1995-02-21
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 13-DEC-1991
APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B. J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 23:

US-09-941-611-23

Query Match	93.5%;	Score 977;	DB 9;	Length 2894;
Best Local Similarity	92.1%;	Pred. NO. 1.2e-76;		
Matches 176;	Conservative	6;	Mismatches 9;	Indels 0;
				Gaps


```
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 179
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
; US-10-651-165-179

Query Match 93.4%; Score 976; DB 15; Length 450;
Best Local Similarity 92.7%; Pred. No. 1.9e-77;
Matches 177; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
Db 1 MSTNPKPKQKTKRNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
QY 61 RRQPIPKARQPGGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGNDPRRRSRNLG 120
Db 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNEGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAALAHGVRATEDGINYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGVAALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 11
US-10-365-620-56
; Sequence 56, Application US/10365620
; Publication No. US20040001853A1
; GENERAL INFORMATION:
; APPLICANT: George, Rajan
; APPLICANT: Tyrrell, Lorne
; APPLICANT: No. US20040001853A1jaim, Antoine
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 656.0016
; CURRENT APPLICATION NUMBER: US/10/365,620
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US60/423,578
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: 60/390,564
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 473
; TYPE: PRT
; ORGANISM: ORF of HCV Core-TBD protein
; US-10-365-620-56

Query Match 93.4%; Score 976; DB 15; Length 473;
Best Local Similarity 92.7%; Pred. No. 2e-77;
Matches 177; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
Db 31 MSTNPKPKQKTKRNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 90
QY 61 RRQPIPKARQPGGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGNDPRRRSRNLG 120
Db 91 RRQPIPKARRPEGRRTWAQPGYPWPLYGNEGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 150
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAALAHGVRATEDGINYATGNLPGCSFSIFLLA 180
Db 151 KVIDTLTCGFADLMGYIPLVGAPLGGVAALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 210
QY 181 LLSCLTTPASA 191
Db 211 LLSCLTTPASA 221

RESULT 12
US-09-742-659-4
; Sequence 4, Application US/09742659
; Patent No. US20010034019A1
; GENERAL INFORMATION:
; APPLICANT: Hong, Zhi
; APPLICANT: Butkiewicz, Nancy J.
; APPLICANT: Zhong, Weidong
; APPLICANT: Ingravallo, Paul
; APPLICANT: Wright-Minogue, Jacquelyn
; APPLICANT: Lau, Johnson Y.
```

; APPLICANT: Lemon, Stanley M.
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
; FILE REFERENCE: ID01116
; CURRENT APPLICATION NUMBER: US/09/742,659
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/171,469
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; US-09-742-659-4

Query Match 93.4%; Score 976; DB 9; Length 3011;
Best Local Similarity 92.7%; Pred. No. 1.5e-76;
Matches 177; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 1 MSTLPKQKTKRNTNRPTDVKFGGGQIVGGVYLLPRRGPRGLGVTRATKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRPPQDVKFGGGQIVGGVYLLPRRGPRGLGVTRATKTSERSQPRG 60
QY 61 RROPIPKARQPGRHWAQPGYPWPLYNCGCGWAGWLLSPRGRPHWGPNDPRRRSRNLG 120
Db 61 RROPIPKARPEGRTWAQPGYPWPLYNCGCGWAGWLLSPRGRSPESWGPTDPRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTVPASA 191

RESULT 13

US-09-952-572-9
; Sequence 9, Application US/09952572
; Patent No. US20020119495A1
; GENERAL INFORMATION:
; APPLICANT: HAWAII BIOTECHNOLOGY GROUP, Inc.
; APPLICANT: NAKANO, Eileen
; APPLICANT: CLEMENTS, David
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITION OF HEPATITIS C
; FILE REFERENCE: HAWBIO1100
; CURRENT APPLICATION NUMBER: US/09/952,572
; CURRENT FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: US 60/230,927
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 9
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
; US-09-952-572-9

Query Match 93.4%; Score 976; DB 9; Length 3011;
Best Local Similarity 92.7%; Pred. No. 1.5e-76;
Matches 177; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 1 MSTLPKQKTKRNTNRPTDVKFGGGQIVGGVYLLPRRGPRGLGVTRATKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRPPQDVKFGGGQIVGGVYLLPRRGPRGLGVTRATKTSERSQPRG 60
QY 61 RROPIPKARQPGRHWAQPGYPWPLYNCGCGWAGWLLSPRGRPHWGPNDPRRRSRNLG 120
Db 61 RROPIPKARPEGRTWAQPGYPWPLYNCGCGWAGWLLSPRGRSPESWGPTDPRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTVPASA 191

RESULT 14

US-09-929-955-1
; Sequence 1, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus sequence
; US-09-929-955-1

Query Match 93.4%; Score 976; DB 9; Length 3011;
Best Local Similarity 92.7%; Pred. No. 1.5e-76;
Matches 177; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 1 MSTLPKQKTKRNTNRPTDVKFGGGQIVGGVYLLPRRGPRGLGVTRATKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRPPQDVKFGGGQIVGGVYLLPRRGPRGLGVTRATKTSERSQPRG 60
QY 61 RROPIPKARQPGRHWAQPGYPWPLYNCGCGWAGWLLSPRGRPHWGPNDPRRRSRNLG 120
Db 61 RROPIPKARPEGRTWAQPGYPWPLYNCGCGWAGWLLSPRGRSPESWGPTDPRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTVPASA 191

RESULT 15

US-09-747-419-20
; Sequence 20, Application US/09747419
; Patent No. US20020155582A1
; GENERAL INFORMATION:
; APPLICANT: Lemon, Stanley
; APPLICANT: Yi, Minkyung
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE
; FILE REFERENCE: 265.0007 0101
; CURRENT APPLICATION NUMBER: US/09/747,419
; CURRENT FILING DATE: 2000-12-23
; PRIOR APPLICATION NUMBER: US 60/171,909
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 20
; LENGTH: 3011

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polyprotein
US-09-747-419-20

Query Match      93.4%; Score 976; DB 9; Length 3011;
Best Local Similarity 92.7%; Pred. No. 1.5e-76;
Matches 177; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSTLPKPKRKTKRNTNRRPTDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPDQVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
QY 61 RRQPIPKARQPGHWAQGYWPWLYGNEGCGWAGWLLSPGSRPHWGPNDRRSRNILG 120
Db 61 RRQPIPKARRPEGRWQAQGYWPWLYGNEGCGWAGWLLSPGSRPSPWGPTDPRRSRNILG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGVARAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGVARAALAHGVRAIEDGVNATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTVFASA 191
```

Search completed: October 30, 2004, 01:49:51
Job time : 130 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2004, 01:46:14 ; Search time 69 Seconds

(without alignments)
993.004 Million cell updates/sec

Title: US-09-084-691B-206

Perfect score: 191

Sequence: 1 MSTLPKQKTKRNTNRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	191	100.0	191	2	AAR92987	Hepatitis
2	51	26.7	97	2	AAR66633	Hepatitis
3	51	26.7	97	2	AAW30597	Hepatitis
4	51	26.7	97	2	AAW01824	Protein e
5	51	26.7	97	2	ADF66080	Hepatitis
6	51	26.7	120	5	AAW50828	Hepatitis
7	51	26.7	170	2	AAR29529	HCV core-
8	51	26.7	190	2	AAR92948	Hepatitis
9	51	26.7	191	2	AAR92949	Hepatitis
10	51	26.7	191	2	AAR92983	Hepatitis
11	51	26.7	191	2	AAR92979	Hepatitis
12	51	26.7	191	2	AAR92955	Hepatitis
13	51	26.7	191	2	AAR92980	Hepatitis
14	51	26.7	191	2	AAR92981	Hepatitis
15	51	26.7	191	2	AAR92942	Hepatitis
16	51	26.7	191	2	AAR92986	Hepatitis
17	51	26.7	191	2	AAR92953	Hepatitis
18	51	26.7	191	2	AAR92972	Hepatitis
19	51	26.7	191	2	AAR92982	Hepatitis
20	51	26.7	191	2	AAR92984	Hepatitis
21	51	26.7	246	2	AAW12714	Flu NS1-H
22	51	26.7	411	2	AAR29541	HCV core-
23	51	26.7	412	2	AAR29539	HCV core-
24	51	26.7	470	2	AAR34470	Encoded b
25	51	26.7	470	2	AAR34472	Encoded b

26	51	26.7	470	2	AAR34478	Encoded b
27	51	26.7	470	2	AAR34474	Encoded b
28	51	26.7	470	2	AAR34477	Encoded b
29	51	26.7	470	2	AAR34476	Encoded b
30	51	26.7	470	2	AAR34479	Encoded b
31	51	26.7	470	2	AAR34473	Encoded b
32	51	26.7	470	2	AAR34475	Encoded b
33	51	26.7	470	2	AAR34469	Encoded b
34	51	26.7	470	2	AAR34467	Encoded b
35	51	26.7	470	2	AAR34471	Hepatitis
36	51	26.7	729	2	AAW47263	Cuticle p
37	51	26.7	729	2	AAW47265	Cuticle p
38	51	26.7	1026	8	ADJ82984	Adenovect
39	51	26.7	1031	2	AAR54067	Non-A, no
40	51	26.7	1031	2	AAR98362	5'UTR/COR
41	51	26.7	3010	5	ABG32458	Hepatitis
42	51	26.7	3010	5	ABG32459	Hepatitis
43	51	26.7	3010	5	ABG32451	Hepatitis
44	51	26.7	3010	5	ABG32455	Hepatitis
45	51	26.7	3010	5	ABG32457	Hepatitis

ALIGNMENTS

RESULT 1
AAR92987
ID AAR92987 standard; protein; 191 AA.

XX AAR92987;
XX
DT 02-OCT-1996 (first entry)
XX
DE Hepatitis C virus isolate HK2 core protein.
XX
KW HCV; EL; envelope 1; core protein; HCV genotyping; antibody; vaccine;
KW hepatitis.
XX
OS Hepatitis C virus.
XX
FN WO9605315-A2.
XX
PD 22-FEB-1996.
XX
PF 15-AUG-1995; 95WO-US010398.
XX
PR 15-AUG-1994; 94US-00290665.
XX (USSH) US SEC DEPT HEALTH.
PA Bukh J, Miller RH, Purcell RH;
PI
XX
DR WPI; 1996-139709/14.
XX N-PSDB; AAT16661.
XX
PT DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
PT determine HCV genotype and as vaccines against HCV infection.
XX
PS Claim 4; Page 223; 340pp; English.
XX
CC AAR92936-R92987 are HCV core proteins derived from 52 different HCV
CC isolates. Isolated cDNA sequences are used for the prodn. of primers
CC useful for detecting the presence of HCV in a sample, the primers are
CC also useful for HCV genotyping. Proteins encoded by the cDNAs can be used
CC in vaccines for immunising against HCV infection. The proteins may also
CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or
CC other mononuclear cells. The antibodies may be used in the prevention of
XX HCV infection
SQ Sequence 191 AA;

Query Match 100.0%; Score 191; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.9e-167;


```
AAV01624
ID AAY01624 standard; protein; 97 AA.
XX
AC AAY01624;
XX
DT 21-JUN-1999 (first entry)
XX
DE Protein encoded by HCV J1 untranslated and core domains.
XX
KW HCV; J7 isolate; J1 isolate; HCV1; immunoassay; asiatic strain;
KW diagnosis; HCV infection; blood screening; immunisation; antiviral.
XX
OS Hepatitis C virus.
XX
PN US5871903-A.
XX
PD 16-FEB-1999.
XX
PF 08-MAY-1995; 95US-00436965.
XX
PR 15-SEP-1989; 89US-00408045.
PR 21-DEC-1989; 89US-00456142.
PR 04-JAN-1991; 91US-00637380.
PR 02-AUG-1993; 93US-00101280.
PR 24-FEB-1994; 94US-00201066.
PR 03-NOV-1994; 94US-00334255.
XX
PA (NAHE-) NAT INST OF HEALTH JAPAN.
PA (CHIR ) CHIRON CORP.
XX
PI Saito I, Miyamura T;
DR WPI; 1999-166619/14.
DR N-PSDB; AAX26743.
XX
PT Immunoassays for Asiatic strains of hepatitis C virus - for diagnosis of
PT infection and screening blood supplies.
XX
PS Disclosure; Fig 18; 43pp; English.
XX
CC The present sequence is encoded by the consensus sequence of the
CC untranslated and core domains of a new hepatitis C virus (HCV), J1. The
CC J1 and J7 (also a new HCV isolate) isolates comprise sequences which are
CC distinct from the prototype HCV isolates, HCV1. The specification
CC describes immunoassays for HCV based on antigens from Asiatic strains not
CC cross-reactive with HCV-1. The assays are used for diagnosis of HCV
CC infection and to screen donated blood. The anti-HCV antibodies are also
CC useful therapeutically and prophylactically (passive immunisation); in
CC screening for antiviral agents; for isolation, purification and
CC identification of non-A, non-B hepatitis virus (e.g. by affinity
CC chromatography) and to raise anti-idiotypic antibodies (useful for
CC treatment or diagnosis and to determine immunogenic regions of the HCV
CC antigens)
XX
SQ Sequence 97 AA;
Query Match 26.7%; Score 51; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 8e-39;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 DVKPPGGQIVGGVYLLPRRGLGVRATKTSERSQPRGRQPIPKARQP 71
Db 21 DVKPPGGQIVGGVYLLPRRGLGVRATKTSERSQPRGRQPIPKARQP 71
RESULT 5
ADP66080
ID ADF66080 standard; protein; 97 AA.
XX
AC ADF66080;
XX
DT 12-FEB-2004 (first entry)
XX
DE Hepatitis C virus J1 domain consensus amino acid sequence.
XX
KW Hepatitis C virus; HCV; screening; identification; non A non B hepatitis;
KW NANBH.
XX
OS Hepatitis C virus.
XX
PN US5959092-A.
XX
PD 28-SEP-1999.
XX
PF 08-MAY-1995; 95US-00436966.
XX
PR 15-SEP-1989; 89US-00408045.
PR 21-DEC-1989; 89US-00456142.
PR 04-JAN-1991; 91US-00637380.
PR 02-AUG-1993; 93US-00101280.
PR 24-FEB-1994; 94US-00201066.
PR 03-NOV-1994; 94US-00334255.
XX
PA (CHIR ) CHIRON CORP.
PA (NAHE-) NAT INST OF HEALTH JAPAN.
XX
PI Saito I, Miyamura T;
DR WPI; 1999-570774/48.
DR N-PSDB; ADF66079.
XX
PT Polynucleotide probes useful for screening and identifying carriers of
PT non A non B hepatitis.
XX
PS Example; Fig 18; 44pp; English.
XX
CC The present invention describes polynucleotides sequences from Hepatitis
CC C virus (HCV) isolates J7 C/E, J1 E, J1 E/NS1, J1 NS3, and J1 NS5 or from
CC deposits BP-2593, BP2594, BP-2595, BP-2637 and BP-2638. The
CC polynucleotide sequences are useful as probes for screening and
CC identifying carriers of non A non B hepatitis (NANBH). The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 97 AA;
Query Match 26.7%; Score 51; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 8e-39;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 DVKPPGGQIVGGVYLLPRRGLGVRATKTSERSQPRGRQPIPKARQP 71
Db 21 DVKPPGGQIVGGVYLLPRRGLGVRATKTSERSQPRGRQPIPKARQP 71
RESULT 6
AAM50828
ID AAM50828 standard; protein; 120 AA.
XX
AC AAM50828;
XX
DT 07-AUG-2003 (revised)
DT 01-MAY-2002 (first entry)
XX
DE Hepatitis C virus core antigen.
XX
KW HCV; core; antigen; immunogen; infection; diagnosis; vaccine.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT Peptide 1..18 /label= MDL-1
FT Peptide 10..24 /label= MDL-2
FT Peptide 28..42 /label= MDL-3
```

FT Peptide 41..57 /label= MDL-4
 FT Peptide 48..65 /label= MDL-5
 FT Peptide 66..80 /label= MDL-6
 FT Peptide 87..101 /label= MDL-7
 FT Peptide 103..120 /label= MDL-8
 XX WO200204484-A2.
 PN 17-JAN-2002.
 PD 06-JUL-2001; 2001WO-CA000988.
 PF 07-JUL-2000; 2000CA-02311022.
 PR (MEDM-) MEDMIRA INC.
 XX Chan HKW, Theolis R;
 PI WPI; 2002-179696/23.
 XX New hepatitis C virus (HCV) peptides and mosaic antigen composition comprising the peptides, useful as immunoreagents for detecting HCV antibodies infection, or as an immunogen for stimulating production of antibodies against HCV.
 XX Disclosure; Page 59; 59pp; English.
 CC The present sequence is that of the core antigen (amino acids 1-120) of the polyprotein encoded by hepatitis C virus (HCV). A panel of overlapping peptides between positions 1 and 186 of the HCV core region was synthesised and tested for immunoreactivity with a commercially available human serum panel. All of the immunoreactivity was concentrated within the region between amino acids 1 to 80. The invention provides a highly immunoreactive mosaic antigen composition (MAC) comprising 2 or more of immunoreactive peptides MDL-1 to MDL-6 (see AAM50816-27) from this region of the core antigen. Each peptide is individually immobilised to a carrier. The unique combination of HCV core peptides in the MAC provides higher specificity and sensitivity for detection of human antibodies specificity to HCV in rapid HCV diagnostic applications. An in vitro diagnostic method for detecting anti-HCV antibodies in a test sample and a diagnostic test kit are provided which use the MAC as an immunoreagent. The HCV MAC may also potentially be used as an immunogen in vaccine compositions. (Updated on 07-AUG-2003 to correct OS field.)
 XX SQ Sequence 120 AA;
 Query Match 26.7%; Score 51; DB 5; Length 120;
 Best Local Similarity 100.0%; Pred. No. 9.5e-39;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 DVKPPGGQIVGGVYLLPRRGLVTRATKTSRSQPRGRROPPIPKARQP 71
 DB 21 DVKPPGGQIVGGVYLLPRRGLVTRATKTSRSQPRGRROPPIPKARQP 71
 RESULT 7
 AAR29529
 ID AAR29529 standard; protein; 170 AA.
 XX AAR29529;
 XX 25-MAR-2003 (revised)
 DT 26-APR-1993 (first entry)
 XX HCV core-envelope peptide N3-1.
 DE Clone; polypeptide; core-envelope; region; Hepatitis C; Virus; HCV; HC;
 KW cDNA; transcriptase; primer; allele; diagnostic; antibody; suppress;

KW control; proteolytic processing; precursor.
 XX Hepatitis C virus.
 XX EP518313-A2.
 XX 16-DEC-1992.
 XX 11-JUN-1992; 92EP-00109812.
 XX 11-JUN-1991; 91JP-00139268.
 PR 12-JUL-1991; 91JP-00172794.
 PR 07-OCT-1991; 91JP-00287008.
 PR 16-DEC-1991; 91JP-00332329.
 PR 20-APR-1992; 92JP-00099957.
 XX (MITU) MITSUBISHI KASEI CORP.
 XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
 PI WPI; 1992-417213/51.
 DR N-PSDB; AAQ32438.
 XX New hepatitis C virus gene and its encoded protein - used for diagnosing
 PT and vaccinating against hepatitis C virus infections.
 XX Claim 5; Page 75-76; 305pp; English.
 CC The sequences given in AAR25929-30 and AAR29534-42 represent various
 CC clones of the polypeptide core-envelope region of the Hepatitis C Virus
 CC (HCV) gene of the invention. These sequences were isolated from the serum
 CC of a patient suffering from hepatitis C (HC). The RNA sequences encoding
 CC these peptides were converted into cDNA using transcriptase in the
 CC presence of one of the primer sequences given in AAQ32540-46. The cDNA
 CC sequences isolated represent different alleles of the same region of the
 CC HCV gene. The entire HCV protein (see AAR29527) is useful in the
 CC development of a diagnostic method which is more accurate and effective
 CC than conventional ones, in the detection of antibodies raised against a
 CC wide range of HCVs which have been hardly detected before. The complete
 CC gene may be used in an in vitro screening system for a substance capable
 CC of specifically suppressing or controlling a proteolytic processing of a
 CC precursor polypeptide of HCV. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX SQ Sequence 170 AA;
 Query Match 26.7%; Score 51; DB 2; Length 170;
 Best Local Similarity 100.0%; Pred. No. 1.3e-38;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 DVKPPGGQIVGGVYLLPRRGLVTRATKTSRSQPRGRROPPIPKARQP 71
 DB 21 DVKPPGGQIVGGVYLLPRRGLVTRATKTSRSQPRGRROPPIPKARQP 71
 RESULT 8
 AAR29248
 ID AAR29248 standard; protein; 190 AA.
 XX AAR29248;
 XX 02-OCT-1996 (first entry)
 DT Hepatitis C virus isolate T10 core protein.
 DE HCV; EI; envelope 1; core protein; HCV genotyping; antibody; vaccine;
 KW hepatitis.
 XX Hepatitis C virus.
 XX WO9605315-A2.
 PN 22-FEB-1996.
 PD


```

RESULT 11
AAR92979
ID AAR92979 standard; protein; 191 AA.
XX AC AAR92979;
XX DT 02-OCT-1996 (first entry)
XX DE Hepatitis C virus isolate SA4 core protein.
XX KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
XX OS Hepatitis C virus.
XX PN WO9605315-A2.
XX PD 22-FEB-1996.
XX PF 15-AUG-1995; 95WO-US010398.
XX PR 15-AUG-1994; 94US-00290665.
XX PA (USSH ) US SEC DEPT HEALTH.
XX PI Bukh J, Miller RH, Purcell RH;
XX DR WPI; 1996-139709/14.
XX DN N-PSDB; AAT16653.
XX PT DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
XX PT determine HCV genotype and as vaccines against HCV infection.
XX PS Claim 4; Page 216-217; 340pp; English.
XX PI Bukh J, Miller RH, Purcell RH;
XX DR WPI; 1996-139709/14.
XX DN N-PSDB; AAT16653.
XX PT DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
XX PT determine HCV genotype and as vaccines against HCV infection.
XX PS Claim 4; Page 216-217; 340pp; English.
XX CC AAR92936-R92987 are HCV core proteins derived from 52 different HCV
XX CC isolates. Isolated cDNA sequences are used for the prodn. of primers
XX CC useful for detecting the presence of HCV in a sample, the primers are
XX CC also useful for HCV genotyping. Proteins encoded by the cDNAs can be used
XX CC in vaccines for immunising against HCV infection. The proteins may also
XX CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or
XX CC other mononuclear cells. The antibodies may be used in the prevention of
XX CC HCV infection
XX SQ Sequence 191 AA;
Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.4e-38;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 DVKPPGGQIVGGVYLLPRGPRGLGVTRATKTSERSQPRGRROPPIPKARQP 71
Db 21 DVKPPGGQIVGGVYLLPRGPRGLGVTRATKTSERSQPRGRROPPIPKARQP 71
RESULT 12
AAR92955
ID AAR92955 standard; protein; 191 AA.
XX AC AAR92955;
XX DT 02-OCT-1996 (first entry)
XX DE Hepatitis C virus isolate HK4 core protein.
XX KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
XX OS Hepatitis C virus.
XX PN WO9605315-A2.
XX PD 22-FEB-1996.
XX PF 15-AUG-1995; 95WO-US010398.
XX PR 15-AUG-1994; 94US-00290665.
XX PA (USSH ) US SEC DEPT HEALTH.
XX PI Bukh J, Miller RH, Purcell RH;
XX DR WPI; 1996-139709/14.
XX DN N-PSDB; AAT16654.
XX PT DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
XX PT determine HCV genotype and as vaccines against HCV infection.
XX PS Claim 4; Page 217; 340pp; English.
XX SQ Sequence 191 AA;
Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.4e-38;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 DVKPPGGQIVGGVYLLPRGPRGLGVTRATKTSERSQPRGRROPPIPKARQP 71
Db 21 DVKPPGGQIVGGVYLLPRGPRGLGVTRATKTSERSQPRGRROPPIPKARQP 71

```

```

PD 22-FEB-1996.
XX 15-AUG-1995; 95WO-US010398.
XX 15-AUG-1994; 94US-00290665.
XX (USSH ) US SEC DEPT HEALTH.
XX PI Bukh J, Miller RH, Purcell RH;
XX DR WPI; 1996-139709/14.
XX DN N-PSDB; AAT16629.
XX PT DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
XX PT determine HCV genotype and as vaccines against HCV infection.
XX PS Claim 4; Page 196-197; 340pp; English.
XX CC AAR92936-R92987 are HCV core proteins derived from 52 different HCV
XX CC isolates. Isolated cDNA sequences are used for the prodn. of primers
XX CC useful for detecting the presence of HCV in a sample, the primers are
XX CC also useful for HCV genotyping. Proteins encoded by the cDNAs can be used
XX CC in vaccines for immunising against HCV infection. The proteins may also
XX CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or
XX CC other mononuclear cells. The antibodies may be used in the prevention of
XX CC HCV infection
XX SQ Sequence 191 AA;
Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.4e-38;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 DVKPPGGQIVGGVYLLPRGPRGLGVTRATKTSERSQPRGRROPPIPKARQP 71
Db 21 DVKPPGGQIVGGVYLLPRGPRGLGVTRATKTSERSQPRGRROPPIPKARQP 71
RESULT 13
AAR92980
ID AAR92980 standard; protein; 191 AA.
XX AC AAR92980;
XX DT 02-OCT-1996 (first entry)
XX DE Hepatitis C virus isolate SA5 core protein.
XX KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
XX KW hepatitis.
XX OS Hepatitis C virus.
XX PN WO9605315-A2.
XX PD 22-FEB-1996.
XX PF 15-AUG-1995; 95WO-US010398.
XX PR 15-AUG-1994; 94US-00290665.
XX PA (USSH ) US SEC DEPT HEALTH.
XX PI Bukh J, Miller RH, Purcell RH;
XX DR WPI; 1996-139709/14.
XX DN N-PSDB; AAT16654.
XX PT DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
XX PT determine HCV genotype and as vaccines against HCV infection.
XX PS Claim 4; Page 217; 340pp; English.
XX SQ Sequence 191 AA;
Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.4e-38;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 DVKPPGGQIVGGVYLLPRGPRGLGVTRATKTSERSQPRGRROPPIPKARQP 71
Db 21 DVKPPGGQIVGGVYLLPRGPRGLGVTRATKTSERSQPRGRROPPIPKARQP 71

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2004, 01:49:56 ; Search time 16 Seconds
(without alignments)
1148.588 Million cell updates/sec

Title: US-09-084-691B-206

Perfect score: 191

Sequence: 1 MSTLPKQKTKRNTNRPT.....CSFSIFLLALLSCLTPASA 191

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	26.7	114	2 S41362	genome polyprotein
2	51	26.7	114	2 S41370	genome polyprotein
3	51	26.7	782	2 S18032	genome polyprotein
4	51	26.7	782	2 S18031	genome polyprotein
5	51	26.7	782	2 S19875	genome polyprotein
6	51	26.7	787	2 P00677	hypothetical prote
7	51	26.7	3010	1 S18030	genome polyprotein
8	50	26.2	112	2 S41371	genome polyprotein
9	50	26.2	114	2 S41365	genome polyprotein
10	49	25.7	82	2 P00804	core protein - hep
11	49	25.7	112	2 S41341	genome polyprotein
12	49	25.7	114	2 S41368	genome polyprotein
13	49	25.7	114	2 S41366	genome polyprotein
14	49	25.7	115	2 S41350	genome polyprotein
15	49	25.7	115	2 S41351	genome polyprotein
16	49	25.7	115	2 S41343	genome polyprotein
17	49	25.7	115	2 S41345	genome polyprotein
18	49	25.7	115	2 S41364	genome polyprotein
19	49	25.7	117	2 S41363	genome polyprotein
20	49	25.7	118	2 S41346	genome polyprotein
21	49	25.7	322	2 J00265	genome polyprotein
22	49	25.7	441	2 S12707	genome polyprotein
23	49	25.7	513	2 A44150	structural protein
24	49	25.7	640	2 J01584	genome polyprotein
25	49	25.7	3010	1 GNWVCJ	genome polyprotein
26	49	25.7	3011	1 GNWVC3	genome polyprotein
27	49	25.7	3011	1 GNWVC3	genome polyprotein
28	47	24.6	88	2 S21336	genome polyprotein
29	47	24.6	108	2 S41357	genome polyprotein

ALIGNMENTS

RESULT 1

S41362

genome polyprotein - hepatitis C virus (genotype 4, N1) (fragment)

N:Contains: core protein

C:Species: hepatitis C virus

A:Variety: genotype 4, N1

C>Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004

C:Accession: S41362

R:van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.

submitted to the EMBL Data Library, January 1994

A:Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A:Reference number: S41341

A:Accession: S41362

A:Molecule type: Genomic RNA

A:Residues: 1-114 <VAN>

A:Cross-references: UNIPROT:O68896; EMBL:Z29465; NID:G443892; PIDN:CAA82603.1; PID:G4438

A:Experimental source: genotype 4, N1

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: capsid protein; core protein; polyprotein

F:1-114/Product: core protein #status predicted <MAT>

Query Match 26.7%; Score 51; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 7.5e-43;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGQIVGGVYLLPRGRLGVRATRTKTSERSQPRGRQPIPKARQP 71

DB 21 DVKPPGGQIVGGVYLLPRGRLGVRATRTKTSERSQPRGRQPIPKARQP 71

RESULT 2

S41370

genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)

N:Contains: core protein

C:Species: hepatitis C virus

A:Variety: genotype 5, N4

C>Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004

C:Accession: S41370

R:van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.

submitted to the EMBL Data Library, January 1994

A:Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A:Reference number: S41341

A:Accession: S41370

A:Molecule type: Genomic RNA

A:Residues: 1-114 <VAN>

A:Cross-references: UNIPROT:O68904; EMBL:Z29473; NID:G443908; PIDN:CAA82611.1; PID:G4439

A:Experimental source: genotype 5, N4

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: capsid protein; core protein; polyprotein

F:1-114/Product: core protein #status predicted <MAT>

Query Match 26.7%; Score 51; DB 2; Length 114;

A:Reference number: S18028
A:Accession: S18030
A:Molecule type: genomic RNA
A:Residues: 1-3010 <HON>
A:Cross-references: UNIPROT:Q68949; EMBL:X61596; NID:g59478; PIDN:CAA43793.1; PID:g59479
A:Experimental source: isolate JK1 from an individual
R:Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A:Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
A:Reference number: A48332; MUID:93119270; PMID:8380322
A:Accession: S33570
A:Molecule type: genomic RNA
A:Residues: 1-547 'T', 549-621 'V', 623-624 'S', 626-652 'DL', 655-761 'T', 763-782 <HON>
A:Cross-references: EMBL:X61591
A:Note: this sequence is inconsistent with the nucleotide translation
A:Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320
as Trp, and TTC for residue 771 as Ser
A:Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBI:P121748)
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F:12-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MER>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus polyprotein NS3 #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4b>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196, 209, 234, 250, 305, 417, 423, 448, 532, 540, 556, 576, 623, 645/Binding site: carbohydrate (As

Query Match 26.7%; Score 51; DB 1; Length 3010;
Best Local Similarity 100.0%; Pred. No. 1.1e-41;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGGQIVGGVLLPRRGPRLGVRATRTKTSERSQPRGRQPIPKARQ 71
Db 21 DVKPPGGGQIVGGVLLPRRGPRLGVRATRTKTSERSQPRGRQPIPKARQ 71

RESULT 8
S41371
genome polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
N:Contains: core protein
C:Species: hepatitis C virus
A:Variety: genotype 5, N5
C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
R:van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A:Reference number: S41341
A:Accession: S41371
A:Molecule type: genomic RNA
A:Residues: 1-112 <VAN>
A:Cross-references: UNIPROT:Q68905; EMBL:Z29474
A:Experimental source: genotype 5, N5
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F:1-112/Product: core protein #status predicted <MAT>

Query Match 26.2%; Score 50; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 7.3e-42;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGGQIVGGVLLPRRGPRLGVRATRTKTSERSQPRGRQPIPKARQ 70
Db 21 DVKPPGGGQIVGGVLLPRRGPRLGVRATRTKTSERSQPRGRQPIPKARQ 70

RESULT 9
S41365
genome polyprotein - hepatitis C virus (genotype 4, N4) (fragment)
N:Contains: core protein
C:Species: hepatitis C virus
A:Variety: genotype 4, N4
C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
R:van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A:Reference number: S41341
A:Accession: S41365
A:Molecule type: genomic RNA
A:Residues: 1-114 <VAN>
A:Cross-references: UNIPROT:Q68899; EMBL:Z29468; NID:g443898; PIDN:CAA82606.1; PID:g44389
A:Experimental source: genotype 4, N4
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F:1-114/Product: core protein #status predicted <MAT>

Query Match 26.2%; Score 50; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 7.4e-42;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGGQIVGGVLLPRRGPRLGVRATRTKTSERSQPRGRQPIPKARQ 70
Db 21 DVKPPGGGQIVGGVLLPRRGPRLGVRATRTKTSERSQPRGRQPIPKARQ 70

RESULT 10
P00804
core protein - hepatitis C virus (fragment)
C:Species: hepatitis C virus
C:Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
R:Simmonds, P.; McOmish, F.; Yap, P.L.; Chan, S.W.; Lin, C.K.; Dusheiko, G.; Saeed, A.A.
J. Gen. Virol. 74, 661-668, 1993
A:Title: Sequence variability in the 5' non-coding region of hepatitis C virus: identification
A:Reference number: P00803; MUID:93224886; PMID:8385694
A:Accession: P00804
A:Molecule type: mRNA
A:Residues: 1-82 <SIM>
A:Cross-references: UNIPROT:Q09740; UNIPROT:Q092971; UNIPROT:Q81564; UNIPROT:Q8JWJ9; UNIP
PROT:Q8QPE3; UNIPROT:Q8JWL6; UNIPROT:Q8V7S7; UNIPROT:Q8V800; UNIPROT:Q8V7Q0; UNIPROT:Q81564; UN
BE2; UNIPROT:Q8B6E1; UNIPROT:Q8B6E0; UNIPROT:Q8B6D7; UNIPROT:Q8B6D5; UN
A:Experimental source: variant EG-33 type 4
A:Accession: P00803
A:Molecule type: mRNA
A:Residues: 1-82 <SI2>
A:Experimental source: variant EG-29 type 4
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein

Query Match 25.7%; Score 49; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 5.5e-41;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGGQIVGGVLLPRRGPRLGVRATRTKTSERSQPRGRQPIPKAR 69
Db 14 DVKPPGGGQIVGGVLLPRRGPRLGVRATRTKTSERSQPRGRQPIPKAR 62

RESULT 11
S41341
genome polyprotein - hepatitis C virus (genotype 1, N1) (fragment)
N:Contains: core protein
C:Species: hepatitis C virus
A:Variety: genotype 1, N1
C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
R:van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994

A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A;Reference number: S41341

A;Accession: S41341

A;Molecule type: genomic RNA

A;Residues: 1-112 <VAN>

A;Cross-references: UNIPROT:Q68875; EMBL:Z29444; NID:G443850; PIDN:CAA82582.1; PID:G4438

A;Experimental source: genotypel, N1

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: capsid protein; core protein; polyprotein

F;1-112/Product: core protein #status predicted <MAT>

Query Match 25.7%; Score 49; DB 2; Length 112;

Best Local Similarity 100.0%; Pred. No. 7.1e-41;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKAR 69

Db 21 DVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKAR 69

RESULT 12

S41368

genome polyprotein - hepatitis C virus (genotype 5, N2) (fragment)

N;Contains: core protein

C;Species: hepatitis C virus

A;Variety: genotype 5, N2

C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004

C;Accession: S41368

R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.

submitted to the EMBL Data Library, January 1994

A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A;Reference number: S41341

A;Accession: S41368

A;Molecule type: genomic RNA

A;Residues: 1-114 <VAN>

A;Cross-references: UNIPROT:Q68902; EMBL:Z29471; NID:G443904; PIDN:CAA82609.1; PID:G4439

A;Experimental source: genotype 5, N2

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: capsid protein; core protein; polyprotein

F;1-114/Product: core protein #status predicted <MAT>

Query Match 25.7%; Score 49; DB 2; Length 114;

Best Local Similarity 100.0%; Pred. No. 7.2e-41;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKAR 69

Db 21 DVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKAR 69

RESULT 13

S41366

genome polyprotein - hepatitis C virus (genotype 4, N5) (fragment)

N;Contains: core protein

C;Species: hepatitis C virus

A;Variety: genotype 4, N5

C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004

C;Accession: S41366

R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.

submitted to the EMBL Data Library, January 1994

A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A;Reference number: S41341

A;Accession: S41366

A;Molecule type: genomic RNA

A;Residues: 1-114 <VAN>

A;Cross-references: UNIPROT:Q68900; EMBL:Z29469; NID:G443900; PIDN:CAA82607.1; PID:G4439

A;Experimental source: genotype 4, N5

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: capsid protein; core protein; polyprotein

F;1-114/Product: core protein #status predicted <MAT>

Query Match 25.7%; Score 49; DB 2; Length 114;

Best Local Similarity 100.0%; Pred. No. 7.2e-41;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKAR 69

Db 21 DVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKAR 69

RESULT 14

S41350

genome polyprotein - hepatitis C virus (genotype 1, N8) (fragment)

N;Contains: core protein

C;Species: hepatitis C virus

A;Variety: genotype 1, N8

C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004

C;Accession: S41350

R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.

submitted to the EMBL Data Library, January 1994

A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A;Reference number: S41341

A;Accession: S41350

A;Molecule type: genomic RNA

A;Residues: 1-115 <VAN>

A;Cross-references: UNIPROT:Q68884; EMBL:Z29453; NID:G443868; PIDN:CAA82591.1; PID:G4438

A;Experimental source: genotype 1, N8

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: capsid protein; core protein; polyprotein

F;1-115/Product: core protein #status predicted <MAT>

Query Match 25.7%; Score 49; DB 2; Length 115;

Best Local Similarity 100.0%; Pred. No. 7.3e-41;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKAR 69

Db 21 DVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKAR 69

RESULT 15

S41351

genome polyprotein - hepatitis C virus (genotype 1, N9) (fragment)

N;Contains: core protein

C;Species: hepatitis C virus

A;Variety: genotype 1, N9

C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004

C;Accession: S41351

R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.

submitted to the EMBL Data Library, January 1994

A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A;Reference number: S41341

A;Accession: S41351

A;Molecule type: genomic RNA

A;Residues: 1-115 <VAN>

A;Cross-references: UNIPROT:Q68885; EMBL:Z29454

A;Experimental source: genotype 1, N9

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: capsid protein; core protein; polyprotein

F;1-115/Product: core protein #status predicted <MAT>

Query Match 25.7%; Score 49; DB 2; Length 115;

Best Local Similarity 100.0%; Pred. No. 7.3e-41;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKAR 69

Db 21 DVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKAR 69

Search completed: October 30, 2004, 01:58:16

Job time : 16 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2004, 01:52:44 ; Search time 67 Seconds
(without alignments)
1640.247 Million cell updates/sec

Title: US-09-084-691B-206
Perfect score: 191
Sequence: 1 MSTLPKPKQKTKRNTNRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 1825181 seqs, 575374646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt 02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	191	100.0	191	2 Q68115	Q68115 hepatitis c
2	161	84.3	191	2 Q68411	Q68411 hepatitis c
3	116	60.7	414	2 Q81270	Q81270 hepatitis c
4	106	55.5	414	2 Q81269	Q81269 hepatitis c
5	97	50.8	117	2 Q75Q30	Q75Q30 hepatitis c
6	97	50.8	117	2 Q75Q31	Q75Q31 hepatitis c
7	97	50.8	117	2 Q75Q33	Q75Q33 hepatitis c
8	97	50.8	117	2 BAD11958	BAD11958 hepatitis c
9	97	50.8	117	2 BAD11960	BAD11960 hepatitis c
10	97	50.8	117	2 BAD11961	BAD11961 hepatitis c
11	86	45.0	414	2 P89956	P89956 hepatitis c
12	85	44.5	414	2 Q81268	Q81268 hepatitis c
13	85	44.5	414	2 Q81329	Q81329 hepatitis c
14	85	44.5	3016	2 Q92531	Q92531 hepatitis c
15	82	42.9	414	2 Q81271	Q81271 hepatitis c
16	81	42.4	414	2 Q81314	Q81314 hepatitis c
17	79	41.4	319	2 Q68703	Q68703 hepatitis c
18	79	41.4	414	2 P89962	P89962 hepatitis c
19	79	41.4	414	2 P89963	P89963 hepatitis c
20	73	38.2	109	2 Q81493	Q81493 hepatitis c
21	67	35.1	158	2 Q80830	Q80830 hepatitis c
22	65	34.0	319	2 Q68690	Q68690 hepatitis c
23	63	33.0	109	2 Q81494	Q81494 hepatitis c
24	63	33.0	415	2 P87751	P87751 hepatitis c
25	63	33.0	3013	2 Q92530	Q92530 hepatitis c
26	56	29.3	3018	2 Q39927	Q39927 hepatitis c
27	55	28.8	319	2 Q68705	Q68705 hepatitis c
28	53	27.7	117	2 Q75Q32	Q75Q32 hepatitis c
29	53	27.7	117	2 BAD11959	BAD11959 hepatitis c
30	53	27.7	414	2 P89958	P89958 hepatitis c
31	51	26.7	74	2 Q68708	Q68708 hepatitis c

32 51 26.7 83 2 Q81264 hepatitis c
33 51 26.7 83 2 Q97980 hepatitis c
34 51 26.7 85 2 Q91KF2 hepatitis c
35 51 26.7 85 2 Q91KF3 hepatitis c
36 51 26.7 85 2 Q91KF4 hepatitis c
37 51 26.7 85 2 Q91KF7 hepatitis c
38 51 26.7 85 2 Q91KG2 hepatitis c
39 51 26.7 85 2 Q91KG3 hepatitis c
40 51 26.7 85 2 Q91KG5 hepatitis c
41 51 26.7 85 2 Q91KG9 hepatitis c
42 51 26.7 85 2 Q91KH0 hepatitis c
43 51 26.7 85 2 Q91KH3 hepatitis c
44 51 26.7 85 2 Q91KH4 hepatitis c
45 51 26.7 85 2 Q91KH5 hepatitis c

ALIGNMENTS

RESULT 1
Q68115 PRELIMINARY; PRT; 191 AA.
AC Q68115; 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Polypeptide (fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94336721; PubMed=8058787;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the core gene of 14 hepatitis C virus
genotypes";
RL Proc. Natl. Acad. Sci. U.S.A. 91:8239-8243(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93376778; PubMed=8396266;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "At least 12 genotypes of hepatitis C virus predicted by sequence
analysis of the putative E1 gene of isolates collected worldwide";
RL Proc. Natl. Acad. Sci. U.S.A. 90:8234-8238(1993).
DR EMBL; U0198; AAA21037.1; -
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; P:Structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Polyprotein.
FT CHAIN 1 >191 core protein.
FT NON TER 191 191
SQ SEQUENCE 191 AA; 20725 MW; 7A7994DF2F0909C6 CRC64;
Query Match 100.0%; Score 191; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.8e-179;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSTLPKPKQKTKRNTNRPTDKPFGGQIYGVYLLPRGPRIGVTRATKTSERSQPRG 60
DB 1 MSTLPKPKQKTKRNTNRPTDKPFGGQIYGVYLLPRGPRIGVTRATKTSERSQPRG 60
QY 61 RROPTPKARQPGGRHWAQGPWPPLYGNCGGACWLLSPRGRPHWGPNDRPRRSRLG 120

```

Db 61 RRQPIPKARQPGHWAQPGYPWPYLYGNECGWAGWLLSPRGRPHWPNDRRRSRNLG 120
QY 121 KVDTLTGCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVDTLTGCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 2
Q68411
ID Q68411 PRELIMINARY; PRT; 191 AA.
AC Q68411; 1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus type 6.
KW Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=42182;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97052554; PubMed=8897188;
RA Barnier L., Willems B., Delage G., Murphy D.G.;
RT "Identification of numerous hepatitis C virus genotypes in Montreal,
RT Canada.";
RL J. Clin. Microbiol. 34:2815-2818 (1996).
DR EMBL; U33435; AAB40038.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
FT NON_TER 191_191
QY SEQUENCE 191 AA; 20789 MW; F292AB64B56DE30A CRC64;

Query Match 84.3%; Score 161; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 6.4e-150;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRGRPIPKARQPGHWAQPG 80
Db 21 DVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRGRPIPKARQPGHWAQPG 80
QY 81 YPWPYLYGNECGWAGWLLSPRGRPHWPNDRRRSRNLGKVIDTLTCGFADLMGYIPV 140
Db 81 YPWPYLYGNECGWAGWLLSPRGRPHWPNDRRRSRNLGKVIDTLTCGFADLMGYIPV 140
QY 141 GAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLAL 191
Db 141 GAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLAL 191

RESULT 3
Q81270
ID Q81270 PRELIMINARY; PRT; 414 AA.
AC Q81270;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
KW Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96005057; PubMed=7561773;
RA Tokita H., Okamoto H., Luengrojanakul P., Vareesangthip K.,

```

```

RA Chainuvati T., Iizuka H., Tsuda F., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Thailand classifiable into five novel
RT genotypes in the sixth (6b), seventh (7c,7d) and ninth (9b, 9c) major
RL J. Gen. Virol. 76:2329-2335 (1995).
DR EMBL; D37845; BAA07091.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002519; HCV core.
DR InterPro; IPR002531; HCV capsid; 1.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 414_414
QY SEQUENCE 414 AA; 44994 MW; 8C1E123FC61D89FC CRC64;

Query Match 60.7%; Score 116; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 2.5e-105;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 WAQGYWPYLYGNECGWAGWLLSPRGRPHWPNDRRRSRNLGKVIDTLTCGFADLMG 135
Db 76 WAQGYWPYLYGNECGWAGWLLSPRGRPHWPNDRRRSRNLGKVIDTLTCGFADLMG 135
QY 136 YIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLALSCLTTPASA 191
Db 136 YIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLALSCLTTPASA 191

RESULT 4
Q81269
ID Q81269 PRELIMINARY; PRT; 414 AA.
AC Q81269;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
KW Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96005057; PubMed=7561773;
RA Tokita H., Okamoto H., Luengrojanakul P., Vareesangthip K.,
RA Chainuvati T., Iizuka H., Tsuda F., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Thailand classifiable into five novel
RT genotypes in the sixth (6b), seventh (7c,7d) and ninth (9b, 9c) major
RL J. Gen. Virol. 76:2329-2335 (1995).
DR EMBL; D37844; BAA07090.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV capsid; 1.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 414_414

```

```
SQ SEQUENCE 414 AA; 45072 MW; 6C9E1AFE4642241D CRC64;
Query Match 55.5%; Score 106; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 1.7e-95;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 WAQPGYPMPLYGNEGCWAGWLLSPRGSRPHWGPNDRPRRSRLGKVIDTLTLCGFADLMG 135
Db 76 WAQPGYPMPLYGNEGCWAGWLLSPRGSRPHWGPNDRPRRSRLGKVIDTLTLCGFADLMG 135

QY 136 YIPVVGAPLGGVAAALAHGVRATEDGINATGNLPCGSFIFLLAL 181
Db 136 YIPVVGAPLGGVAAALAHGVRATEDGINATGNLPCGSFIFLLAL 181

RESULT 5
Q75Q30 PRELIMINARY; PRT; 117 AA.
AC Q75Q30;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Polyprotein (fragment).
OS Hepatitis C virus type 6a.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus; Hepatitis C virus type 6.
OX NCBI_TaxID=31655;
RN [1]_
RP SEQUENCE FROM N.A.
RA Hirano M., Tran H.T., Abe K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB162867; BAD11961.1; -.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER 117
SQ SEQUENCE 117 AA; 13441 MW; 9EA99F0E118A3F20 CRC64;

Query Match 50.8%; Score 97; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.6e-87;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFFGGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRGRQPIPKARQPOGRHWAQPG 80
Db 21 DVKFFGGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRGRQPIPKARQPOGRHWAQPG 80

QY 81 YPWPLYGNEGCWAGWLLSPRGSRPHWGPNDRPRRSR 117
Db 81 YPWPLYGNEGCWAGWLLSPRGSRPHWGPNDRPRRSR 117

RESULT 6
Q75Q31 PRELIMINARY; PRT; 117 AA.
AC Q75Q31;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Polyprotein (fragment).
OS Hepatitis C virus type 6a.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus; Hepatitis C virus type 6.
OX NCBI_TaxID=31655;
RN [1]_
RP SEQUENCE FROM N.A.
RA Hirano M., Tran H.T., Abe K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB162866; BAD11960.1; -.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER 117
SQ SEQUENCE 117 AA; 13407 MW; 3EAE29D8CAE789F6 CRC64;

Query Match 50.8%; Score 97; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.6e-87;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFFGGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRGRQPIPKARQPOGRHWAQPG 80
Db 21 DVKFFGGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRGRQPIPKARQPOGRHWAQPG 80

QY 81 YPWPLYGNEGCWAGWLLSPRGSRPHWGPNDRPRRSR 117
Db 81 YPWPLYGNEGCWAGWLLSPRGSRPHWGPNDRPRRSR 117

RESULT 7
Q75Q33 PRELIMINARY; PRT; 117 AA.
AC Q75Q33;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Polyprotein (fragment).
OS Hepatitis C virus type 6a.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus; Hepatitis C virus type 6.
OX NCBI_TaxID=31655;
RN [1]_
RP SEQUENCE FROM N.A.
RA Hirano M., Tran H.T., Abe K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB162864; BAD11958.1; -.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER 117
SQ SEQUENCE 117 AA; 13407 MW; 3EAE29D8CAE789F6 CRC64;

Query Match 50.8%; Score 97; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.6e-87;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFFGGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRGRQPIPKARQPOGRHWAQPG 80
Db 21 DVKFFGGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRGRQPIPKARQPOGRHWAQPG 80

QY 81 YPWPLYGNEGCWAGWLLSPRGSRPHWGPNDRPRRSR 117
Db 81 YPWPLYGNEGCWAGWLLSPRGSRPHWGPNDRPRRSR 117

RESULT 8
BAD11958 PRELIMINARY; PRT; 117 AA.
AC BAD11958;
DT 03-MAR-2004 (TrEMBLrel. 27, Created)
DT 03-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE Polyprotein (fragment).
OS Hepatitis C virus type 6a.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus; Hepatitis C virus; Hepatitis C virus type 6.
OX NCBI_TaxID=31655;
RN [1]_
RP SEQUENCE FROM N.A.
RA Hirano M., Tran H.T., Abe K.;
RT "Genotypic distribution of hepatitis C virus (HCV) in Ho Chi Minh City, Vietnam: New genotyping systems for identification of Vietnamese HCV isolates.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB162864; BAD11958.1; -.
KW Polyprotein.
FT NON_TER 117
```

SQ	SEQUENCE	117 AA; 13407 MW; 3EAE29D8CAE789F6 CRC64;
Query Match	50.8%; Score 97; DB 2; Length 117;	
Best Local Similarity	100.0%; Pred. No. 4.6e-87;	
Matches	97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	21 DVKFPGGQIVGGVYLLPRGPRILGVTRATKTSERSQPRGRORIPKAROPQGRHWAQPG 80	
DB	21 DVKFPGGQIVGGVYLLPRGPRILGVTRATKTSERSQPRGRORIPKAROPQGRHWAQPG 80	
QY	81 YPWPLYNCGCGWAGWLLSPRGSRPHWGPNDPRRSR 117	
DB	81 YPWPLYNCGCGWAGWLLSPRGSRPHWGPNDPRRSR 117	
RESULT 9		
BAD11960		
ID	BAD11960 PRELIMINARY; PRT; 117 AA.	
AC	BAD11960;	
DT	03-MAR-2004 (TREMBlrel. 27, Created)	
DT	03-MAR-2004 (TREMBlrel. 27, Last sequence update)	
DT	03-MAR-2004 (TREMBlrel. 27, Last annotation update)	
DE	Polyprotein (Fragment).	
OS	Hepatitis C virus type 6a.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	
OC	Hepacivirus; Hepatitis C virus; Hepatitis C virus type 6.	
OX	NCBI_TaxID=31655;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=VT657;	
RA	Hirano M., Tran H.T., Abe K.;	
RT	"Genotypic distribution of hepatitis C virus (HCV) in Ho Chi Minh City, Vietnam: New genotyping systems for identification of Vietnamese HCV isolates.";	
RT	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.	
RL	EMBL; AB162866; BAD11960.1; -.	
KW	Polyprotein.	
FT	NON_TER 117 117	
SQ	SEQUENCE 117 AA; 13407 MW; 3EAE29D8CAE789F6 CRC64;	
Query Match	50.8%; Score 97; DB 2; Length 117;	
Best Local Similarity	100.0%; Pred. No. 4.6e-87;	
Matches	97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	21 DVKFPGGQIVGGVYLLPRGPRILGVTRATKTSERSQPRGRORIPKAROPQGRHWAQPG 80	
DB	21 DVKFPGGQIVGGVYLLPRGPRILGVTRATKTSERSQPRGRORIPKAROPQGRHWAQPG 80	
QY	81 YPWPLYNCGCGWAGWLLSPRGSRPHWGPNDPRRSR 117	
DB	81 YPWPLYNCGCGWAGWLLSPRGSRPHWGPNDPRRSR 117	
RESULT 10		
BAD11961		
ID	BAD11961 PRELIMINARY; PRT; 117 AA.	
AC	BAD11961;	
DT	03-MAR-2004 (TREMBlrel. 27, Created)	
DT	03-MAR-2004 (TREMBlrel. 27, Last sequence update)	
DT	03-MAR-2004 (TREMBlrel. 27, Last annotation update)	
DE	Polyprotein (Fragment).	
OS	Hepatitis C virus type 6a.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	
OC	Hepacivirus; Hepatitis C virus; Hepatitis C virus type 6.	
OX	NCBI_TaxID=31655;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=VT689;	
RA	Hirano M., Tran H.T., Abe K.;	
RT	"Genotypic distribution of hepatitis C virus (HCV) in Ho Chi Minh City, Vietnam: New genotyping systems for identification of Vietnamese HCV isolates.";	
RT	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.	
RL	EMBL; AB162867; BAD11961.1; -.	
KW	Polyprotein.	
FT	NON_TER 117 117	
SQ	SEQUENCE 117 AA; 13441 MW; 9EA99F0E118A3F20 CRC64;	
Query Match	50.8%; Score 97; DB 2; Length 117;	
Best Local Similarity	100.0%; Pred. No. 4.6e-87;	
Matches	97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	21 DVKFPGGQIVGGVYLLPRGPRILGVTRATKTSERSQPRGRORIPKAROPQGRHWAQPG 80	
DB	21 DVKFPGGQIVGGVYLLPRGPRILGVTRATKTSERSQPRGRORIPKAROPQGRHWAQPG 80	
QY	81 YPWPLYNCGCGWAGWLLSPRGSRPHWGPNDPRRSR 117	
DB	81 YPWPLYNCGCGWAGWLLSPRGSRPHWGPNDPRRSR 117	
RESULT 11		
P89956		
ID	P89956 PRELIMINARY; PRT; 414 AA.	
AC	P89956;	
DT	01-MAY-1997 (TREMBlrel. 03, Created)	
DT	01-MAY-1997 (TREMBlrel. 03, Last sequence update)	
DT	01-OCT-2003 (TREMBlrel. 25, Last annotation update)	
DE	Core, env and part of E2/NS1 (Fragment).	
OS	Hepatitis C virus.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	
OC	Hepacivirus.	
OX	NCBI_TaxID=11103;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
EX	MEDLINE=95062197; PubMed=7972001;	
RA	Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,	
RA	Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.;	
RT	"Hepatitis C virus variants from Vietnam are classifiable into the seventh, eighth, and ninth major genetic groups.";	
RT	Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).	
RL	[2]	
RP	SEQUENCE FROM N.A.	
RA	Okamoto H.;	
RL	Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.	
EMBL	D88469; BAA13618.1; -.	
PIR	PQ0804; PQ0804.	
DR	GO; GO:0016021; C:integral to membrane; IEA.	
DR	GO; GO:0019028; C:viral capsid; IEA.	
DR	GO; GO:0019031; C:viral envelope; IEA.	
DR	GO; GO:0005198; F:structural molecule activity; IEA.	
DR	InterPro; IPR002522; HCV capsid.	
DR	InterPro; IPR002521; HCV core.	
DR	InterPro; IPR002519; HCV env.	
DR	InterPro; IPR002531; HCV NS1.	
DR	Pfam; PF01543; HCV capsid; 1.	
DR	Pfam; PF01542; HCV core; 1.	
DR	Pfam; PF01539; HCV env; 1.	
DR	Pfam; PF01560; HCV NS1; 1.	
KW	Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;	
FT	NON_TER 414 414	
SQ	SEQUENCE 414 AA; 44528 MW; 7A3BE1710311C017 CRC64;	
Query Match	45.0%; Score 86; DB 2; Length 414;	
Best Local Similarity	100.0%; Pred. No. 8.6e-76;	
Matches	86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	72 QGRHWAQGYWPPLYNCGCGWAGWLLSPRGSRPHWGPNDPRRSRNLGKVIDTLTCGFA 131	
DB	72 QGRHWAQGYWPPLYNCGCGWAGWLLSPRGSRPHWGPNDPRRSRNLGKVIDTLTCGFA 131	
QY	132 DLMGYIPVVGAPLGGVAAALAHGVA 157	
DB	132 DLMGYIPVVGAPLGGVAAALAHGVA 157	


```
RESULT 12
ID Q81268 PRELIMINARY; PRT; 414 AA.
AC Q81268;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Core, env, and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96005057; PubMed=7561773;
Tokita H., Okamoto H., Luengrojjanakul P., Vareesangthip K.,
Chaiyavutti T., Iizuka H., Tsuda F., Miyakawa Y., Mayumi M.;
"Hepatitis C virus variants from Thailand classifiable into five novel
genotypes in the sixth (6b), seventh (7c,7d) and ninth (9b, 9c) major
genetic groups.";
J. Gen. Virol. 76:2329-2335 (1995).
RL EMBL; D37843; BAA07089.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002531; HCV_NS1.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01560; HCV_NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
FT NON_TER 414
SQ SEQUENCE 414 AA; 44836 MW; 212740491A9DA0B1 CRC64;

Query Match 44.5%; Score 85; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 8.4e-75;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 WGNPDRRRNLGKVIDITLCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVAT 166
Db 107 WGNPDRRRNLGKVIDITLCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVAT 166
QY 167 GNLPCCSFIFLLALLSCLITTPASA 191
Db 167 GNLPCCSFIFLLALLSCLITTPASA 191

RESULT 14
ID Q92531 PRELIMINARY; PRT; 3016 AA.
AC Q92531;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98378034; PubMed=9714232;
Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda F., Miyakawa Y.,
Mayumi M.;
"The entire nucleotide sequences of three hepatitis C virus isolates
in genetic groups 7-9 and comparison with those in the other eight
genetic groups.";
J. Gen. Virol. 79:1847-1857 (1998).
RL EMBL; D84264; BAA32666.1; -.
DR PIR; PQ0804; PQ0804.
DR HSSP; P27958; 1A1V.
DR MEROPS; S29.001; -.
DR MEROPS; U39.001; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.

QY 107 WGNPDRRRNLGKVIDITLCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVAT 166
Db 107 WGNPDRRRNLGKVIDITLCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINVAT 166
QY 167 GNLPCCSFIFLLALLSCLITTPASA 191
Db 167 GNLPCCSFIFLLALLSCLITTPASA 191

RESULT 13
ID Q81329 PRELIMINARY; PRT; 414 AA.
AC Q81329;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;
Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.;
"Hepatitis C virus variants from Vietnam are classifiable into the
seventh, eighth, and ninth major genetic groups.";
Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026 (1994).
RL
```

```
DR GO:0019079; P:Viral genome replication; IEA.
DR GO:0019087; P:Viral transformation; IEA.
DR InterPro; IPR00345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SMO0487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3016 AA; 328032 MW; 4E5CFF96258BCE3B CRC64;

Query Match 44.5%; Score 85; DB 2; Length 3016;
Best Local Similarity 100.0%; Pred. No. 4.le-74;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 WGNPDRRRNRNLGKVIDTLTCGFADLMGYPVVGAPLGVAALAHGVRAIEDGINYAT 166
DB 107 WGNPDRRRNRNLGKVIDTLTCGFADLMGYPVVGAPLGVAALAHGVRAIEDGINYAT 166

QY 167 GNLPGCSFISFLALLSLCTTPASA 191
DB 167 GNLPGCSFISFLALLSLCTTPASA 191

RESULT 15
Q81271 ID Q81271 PRELIMINARY; PRT; 414 AA.
AC Q81271;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96005057; PubMed=7561773;
RA Tokita H., Okamoto H., Luengrojjanakul P., Varesangthip K.,
RA Chainuvati T., Lizuka H., Tsuda F., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Thailand classifiable into five novel
RT genotypes in the sixth (6b), seventh (7c,7d) and ninth (9b, 9c) major
RT genetic groups.";
RL J. Gen. Virol. 76:2329-2335(1995).
DR EMBL; D37846; BAA07092.1;
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0019028; C:viral capsid; IEA.
DR GO:0019031; C:viral envelope; IEA.
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
```

```
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV_NS1.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 414 414
SQ SEQUENCE 414 AA; 45148 MW; D63EE7CED5B71776 CRC64;

Query Match 42.9%; Score 82; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 7.5e-72;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 WAQPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGNPDRRRSRNLGKVIDTLTCGFADLMG 135
DB 76 WAQPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGNPDRRRSRNLGKVIDTLTCGFADLMG 135

QY 136 YIPVVGAPLGVAALAHGVRA 157
DB 136 YIPVVGAPLGVAALAHGVRA 157

Search completed: October 30, 2004, 01:59:29
Job time : 68 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2004, 01:57:19 ; Search time 21 Seconds
(without alignments)
603.178 Million cell updates/sec

Title: US-09-084-691B-206

Perfect score: 191

Sequence: 1 MSTLPKPKQKTKNTNRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 478139 seqs, 66318000 residues

Word size : 0

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2.6/prodata/1/iaa/5A COMB.pep.*
- 2: /cgn2.6/prodata/1/iaa/5B COMB.pep.*
- 3: /cgn2.6/prodata/1/iaa/6A COMB.pep.*
- 4: /cgn2.6/prodata/1/iaa/6B COMB.pep.*
- 5: /cgn2.6/prodata/1/iaa/6C COMB.pep.*
- 6: /cgn2.6/prodata/1/iaa/6D COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	191	100.0	191	2	US-08-290-665A-206
2	191	100.0	191	5	PCT-US95-10398-206
3	191	100.0	319	4	US-08-635-886C-228
4	191	100.0	319	4	US-08-974-690C-228
5	51	26.7	191	2	US-08-290-665A-161
6	51	26.7	191	2	US-08-290-665A-167
7	51	26.7	191	2	US-08-290-665A-168
8	51	26.7	191	2	US-08-290-665A-172
9	51	26.7	191	2	US-08-290-665A-174
10	51	26.7	191	2	US-08-290-665A-191
11	51	26.7	191	2	US-08-290-665A-198
12	51	26.7	191	2	US-08-290-665A-199
13	51	26.7	191	2	US-08-290-665A-200
14	51	26.7	191	2	US-08-290-665A-201
15	51	26.7	191	2	US-08-290-665A-202
16	51	26.7	191	2	US-08-290-665A-203
17	51	26.7	191	2	US-08-290-665A-205
18	51	26.7	191	5	PCT-US95-10398-161
19	51	26.7	191	5	PCT-US95-10398-167
20	51	26.7	191	5	PCT-US95-10398-168
21	51	26.7	191	5	PCT-US95-10398-172
22	51	26.7	191	5	PCT-US95-10398-174
23	51	26.7	191	5	PCT-US95-10398-191
24	51	26.7	191	5	PCT-US95-10398-198
25	51	26.7	191	5	PCT-US95-10398-199
26	51	26.7	191	5	PCT-US95-10398-200
27	51	26.7	191	5	PCT-US95-10398-201

28	51	26.7	191	5	PCT-US95-10398-202	Sequence 202, App
29	51	26.7	191	5	PCT-US95-10398-203	Sequence 203, App
30	51	26.7	191	5	PCT-US95-10398-205	Sequence 205, App
31	51	26.7	450	4	US-08-635-886C-189	Sequence 189, App
32	51	26.7	450	4	US-08-635-886C-194	Sequence 194, App
33	51	26.7	450	4	US-08-974-690C-189	Sequence 189, App
34	51	26.7	450	4	US-08-974-690C-194	Sequence 194, App
35	51	26.7	3010	4	US-09-539-601-3	Sequence 3, Appli
36	51	26.7	3010	4	US-09-539-601-21	Sequence 21, Appl
37	51	26.7	3010	4	US-09-539-601-27	Sequence 27, Appl
38	51	26.7	3010	4	US-09-539-601-33	Sequence 33, Appl
39	50	26.2	191	2	US-08-290-665A-197	Sequence 197, App
40	50	26.2	191	5	PCT-US95-10398-197	Sequence 197, App
41	50	26.2	319	4	US-08-635-886C-217	Sequence 217, App
42	50	26.2	319	4	US-08-974-690C-217	Sequence 217, App
43	49	25.7	75	4	US-08-905-054B-1	Sequence 1, Appli
44	49	25.7	75	5	PCT-US92-07813-1	Sequence 1, Appli
45	49	25.7	90	1	US-07-681-703B-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-290-665A-206
; Sequence 206, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 206:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: hom sapiens
; INDIVIDUAL ISOLATE: HK2
US-08-290-665A-206

Query Match 100.0%; Score 191; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1e-173;

Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKNTNRPTDVKPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
 Db 1 MSTLPKPKQKTKNTNRPTDVKPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60

QY 61 RRQIPKARQPGQRHWAQPGYPWPLYGNECGWAGWLLSPRGRPHWGPNDPRRRSRNLG 120
 Db 61 RRQIPKARQPGQRHWAQPGYPWPLYGNECGWAGWLLSPRGRPHWGPNDPRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
 Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
 Db 181 LLSCLTTPASA 191

RESULT 2
 PCT-US95-10398-206
 ; Sequence 206, Application PC/TUS9510398
 ; GENERAL INFORMATION:
 ; APPLICANT: BUKH, J., MILLER, R.H. AND
 ; APPLICANT: PURCELL, R.H.
 ; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
 ; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
 ; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
 ; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
 ; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
 ; NUMBER OF SEQUENCES: 263
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORGAN & FINNEGAN
 ; STREET: 345 PARK AVENUE
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: USA
 ; ZIP: 10154
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WORDPERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/10398
 ; FILING DATE: 15-AUG-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/086,428
 ; FILING DATE: 29 JUNE 1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/290/665
 ; FILING DATE: 15 AUGUST 1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: RICHARD W. BORK
 ; REGISTRATION NUMBER: 36,459
 ; REFERENCE/DOCKET NUMBER: 2026-4116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 758-4800
 ; TELEFAX: (212) 751-6849
 ; TELEX: 421792
 ; INFORMATION FOR SEQ ID NO: 206:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 191 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; ORGANISM: homopisapiens
 ; INDIVIDUAL ISOLATE: HK2
 ; PCT-US95-10398-206

Query Match 100.0%; Score 191; DB 5; Length 191;

Best Local Similarity 100.0%; Pred. No. 1e-173;
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKNTNRPTDVKPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
 Db 1 MSTLPKPKQKTKNTNRPTDVKPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60

QY 61 RRQIPKARQPGQRHWAQPGYPWPLYGNECGWAGWLLSPRGRPHWGPNDPRRRSRNLG 120
 Db 61 RRQIPKARQPGQRHWAQPGYPWPLYGNECGWAGWLLSPRGRPHWGPNDPRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
 Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
 Db 181 LLSCLTTPASA 191

RESULT 3
 US-08-635-886C-228
 ; Sequence 228, Application US/08635886C
 ; Patent No. 6555114
 ; GENERAL INFORMATION:
 ; APPLICANT: LEROUX-ROELS, Geert
 ; APPLICANT: DELEYS, Robert
 ; APPLICANT: MAERTENS, Geert
 ; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
 ; TITLE OF INVENTION: VIRUS
 ; FILE REFERENCE: 2752-18
 ; CURRENT APPLICATION NUMBER: US/08/635,886C
 ; CURRENT FILING DATE: 1996-04-25
 ; PRIOR APPLICATION NUMBER: PCT/EP94/03555
 ; PRIOR FILING DATE: 1994-10-28
 ; PRIOR APPLICATION NUMBER: EP 93402718.6
 ; PRIOR FILING DATE: 1993-11-04
 ; NUMBER OF SEQ ID NOS: 286
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 228
 ; LENGTH: 319
 ; TYPE: PRT
 ; ORGANISM: hepatitis C virus
 ; US-08-635-886C-228

Query Match 100.0%; Score 191; DB 4; Length 319;
 Best Local Similarity 100.0%; Pred. No. 1.6e-173;
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKNTNRPTDVKPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
 Db 1 MSTLPKPKQKTKNTNRPTDVKPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60

QY 61 RRQIPKARQPGQRHWAQPGYPWPLYGNECGWAGWLLSPRGRPHWGPNDPRRRSRNLG 120
 Db 61 RRQIPKARQPGQRHWAQPGYPWPLYGNECGWAGWLLSPRGRPHWGPNDPRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
 Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
 Db 181 LLSCLTTPASA 191

RESULT 4
 US-08-974-690C-228
 ; Sequence 228, Application US/08974690C
 ; Patent No. 6613333
 ; GENERAL INFORMATION:
 ; APPLICANT: LEROUX-ROELS, Geert
 ; APPLICANT: DELEYS, Robert

APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
FILE REFERENCE: 2551-94
CURRENT APPLICATION NUMBER: US/08/974,690C
PRIOR FILING DATE: 1997-11-19
PRIOR APPLICATION NUMBER: PCT/EP94/03555
PRIOR FILING DATE: 1994-10-28
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: Patentin version 3.1
SEQ ID NO 228
LENGTH: 319
TYPE: PRT
ORGANISM: hepatitis C virus
US-08-974-690C-228

Query Match 100.0%; Score 191; DB 4; Length 319;
Best Local Similarity 100.0%; Pred. No. 1.6e-173;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFPGGGQIVGGVYLLPRGPRGPRGPRGPRGPRGPRG 60
DB 1 MSTLPKPKQKTKRNTNRRPTDVKFPGGGQIVGGVYLLPRGPRGPRGPRGPRGPRGPRG 60
QY 61 RQPIPKARQPGRHWAGQYWPPLYGNEGCGWAGWLLSPRGRPHWPNDRRRSRNLG 120
DB 61 RQPIPKARQPGRHWAGQYWPPLYGNEGCGWAGWLLSPRGRPHWPNDRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGNGYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGNGYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 5
US-08-290-665A-161
Sequence 161, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 167:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: T10
US-08-290-665A-167

Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e-40;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFGGGQIVGGVYLLPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRG 71
DB 21 DVKFGGGQIVGGVYLLPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRG 71

RESULT 6
US-08-290-665A-167
Sequence 167, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 167:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: T10
US-08-290-665A-167

Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e-40;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKARQP 71
|||||
Db 21 DVKPPGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKARQP 71
|||||

RESULT 7

US-08-290-665A-168
; Sequence 168, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 168:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SW2
US-08-290-665A-168

Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e-40;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKARQP 71
|||||
Db 21 DVKPPGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKARQP 71
|||||

RESULT 8

US-08-290-665A-172
; Sequence 172, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND

; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 172:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: HK3
US-08-290-665A-172

Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e-40;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKARQP 71
|||||
Db 21 DVKPPGGQIVGGVYLLPRGPRGLGVRATKTSRSQPRGRQPIPKARQP 71
|||||

RESULT 9

US-08-290-665A-174
; Sequence 174, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1

RESULT 12
US-08-290-665A-199
; Sequence 199, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BURKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 199:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORGANISM: homosapiens
; ORIGINAL SOURCE: SA5
; INDIVIDUAL ISOLATE: SA5
US-08-290-665A-199
Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e-40;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 DVKPPGGQIVGGVYLLPRGRLGVRATKTSERSQPRGRQPIPKARQP 71
Db 21 DVKPPGGQIVGGVYLLPRGRLGVRATKTSERSQPRGRQPIPKARQP 71
RESULT 13
US-08-290-665A-200
; Sequence 200, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BURKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE

; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 200:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORGANISM: homosapiens
; ORIGINAL SOURCE: SA7
; INDIVIDUAL ISOLATE: SA7
US-08-290-665A-200
Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e-40;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 DVKPPGGQIVGGVYLLPRGRLGVRATKTSERSQPRGRQPIPKARQP 71
Db 21 DVKPPGGQIVGGVYLLPRGRLGVRATKTSERSQPRGRQPIPKARQP 71
RESULT 14
US-08-290-665A-201
; Sequence 201, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BURKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459

REFERENCE/DOCKET NUMBER: 2026-4116
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 201:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: SA1
US-08-290-665A-201

Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e-40;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGGQIVGGVILLPRRGLGVRAIRKTSERSQPRGRQPIPKARQP 71
DB 21 DVKFPGGGQIVGGVILLPRRGLGVRAIRKTSERSQPRGRQPIPKARQP 71

RESULT 15

US-08-290-665A-202
Sequence 202, Application US/08290665A
Patent No. 5882852

GENERAL INFORMATION:

APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.

TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792

INFORMATION FOR SEQ ID NO: 202:

SEQUENCE CHARACTERISTICS:

LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:

ORGANISM: homosapiens
INDIVIDUAL ISOLATE: SA3

US-08-290-665A-202

Query Match 26.7%; Score 51; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e-40;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGGQIVGGVILLPRRGLGVRAIRKTSERSQPRGRQPIPKARQP 71
DB 21 DVKFPGGGQIVGGVILLPRRGLGVRAIRKTSERSQPRGRQPIPKARQP 71

Search completed: October 30, 2004, 02:00:42
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2004, 01:58:21 ; Search time 49 Seconds

(without alignments)
1263.780 Million cell updates/sec

Title: US-09-084-691B-206

Perfect score: 191

Sequence: 1 MSTLPKPKQKTKRNTNRPT.....CSFSIFLLALLSCLTPASA 191

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1370721 seqs, 324215800 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	100.0	319	15	US-10-651-165-228
2	51	26.7	450	15	US-10-651-165-189
3	51	26.7	450	15	US-10-651-165-194
4	51	26.7	3010	15	US-10-467-000-1
5	50	26.2	319	15	US-10-651-165-217
6	49	25.7	75	15	US-10-431-587-1
7	49	25.7	77	9	US-09-921-397-3
8	49	25.7	91	9	US-09-758-308-1
9	49	25.7	94	10	US-09-891-983A-16
10	49	25.7	94	14	US-10-173-480-16
11	49	25.7	94	16	US-10-753-910-16
12	49	25.7	96	10	US-09-899-046-192
13	49	25.7	96	10	US-09-878-281-192

14 49 25.7 96 10 US-09-873-224-192 Sequence 192, App
15 49 25.7 97 9 US-09-756-875-8 Sequence 8, Appli
16 49 25.7 103 9 US-09-921-397-77 Sequence 77, Appl
17 49 25.7 113 9 US-09-921-397-78 Sequence 78, Appl
18 49 25.7 117 9 US-09-851-138-28 Sequence 28, Appl
19 49 25.7 117 15 US-10-651-165-225 Sequence 225, App
20 49 25.7 120 9 US-09-306-780-4 Sequence 4, Appli
21 49 25.7 122 14 US-10-098-857B-1 Sequence 1, Appli
22 49 25.7 126 10 US-09-899-046-166 Sequence 166, App
23 49 25.7 126 10 US-09-878-281-166 Sequence 166, App
24 49 25.7 126 10 US-09-873-224-166 Sequence 166, App
25 49 25.7 130 14 US-10-268-569-19 Sequence 19, Appl
26 49 25.7 150 9 US-09-306-780-6 Sequence 6, Appli
27 49 25.7 151 14 US-10-292-129-14 Sequence 14, Appl
28 49 25.7 158 9 US-09-851-138-66 Sequence 66, Appli
29 49 25.7 161 9 US-09-306-780-8 Sequence 8, Appli
30 49 25.7 166 10 US-09-899-046-164 Sequence 164, App
31 49 25.7 166 10 US-09-899-046-194 Sequence 194, App
32 49 25.7 166 10 US-09-878-281-164 Sequence 164, App
33 49 25.7 166 10 US-09-878-281-194 Sequence 194, App
34 49 25.7 166 10 US-09-873-224-164 Sequence 164, App
35 49 25.7 166 10 US-09-873-224-194 Sequence 194, App
36 49 25.7 176 9 US-09-306-780-20 Sequence 20, Appl
37 49 25.7 182 9 US-09-929-955-2 Sequence 2, Appli
38 49 25.7 182 13 US-10-104-966-2 Sequence 2, Appli
39 49 25.7 182 15 US-10-719-619-2 Sequence 2, Appli
40 49 25.7 190 14 US-10-268-562-1 Sequence 1, Appli
41 49 25.7 190 15 US-10-450-649-7 Sequence 7, Appli
42 49 25.7 191 9 US-09-306-780-10 Sequence 10, Appl
43 49 25.7 191 10 US-09-194-949-3 Sequence 3, Appli
44 49 25.7 235 15 US-10-365-620-58 Sequence 58, Appl
45 49 25.7 249 15 US-10-365-620-54 Sequence 54, Appl

ALIGNMENTS

RESULT 1

US-10-651-165-228
; Sequence 228, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 228
; LENGTH: 319
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-228

Query Match 100.0%; Score 191; DB 15; Length 319;
Best Local Similarity 100.0%; Pred. No. 6.5e-166;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSTLPKPKQKTKRNTNRPTDKPFGGQIVGGVYLLPRRGLGVRATRKTSRSQPRG 60
Db 1 MSTLPKPKQKTKRNTNRPTDKPFGGQIVGGVYLLPRRGLGVRATRKTSRSQPRG 60
QY 61 RQPIPKARQPKRHWAPGYWPFLYGNCGCWAGWLLSPRGRPHWGPNDPRRSRLG 120

Db 61 RRPPIKARQPGHQAQPGYFPLVNGECCHAGWLLSPRGSRPHWGNDRPRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 2
US-10-651-165-189
; Sequence 189, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651.165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 189
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-189

Query Match 26.7%; Score 51; DB 15; Length 450;
Best Local Similarity 100.0%; Pred. No. 5.1e-38;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRQPIPKARQP 71
Db 21 DVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRQPIPKARQP 71

RESULT 3
US-10-651-165-194
; Sequence 194, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651.165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 194
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus

US-10-651-165-194
Query Match 26.7%; Score 51; DB 15; Length 450;
Best Local Similarity 100.0%; Pred. No. 5.1e-38;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRQPIPKARQP 71
Db 21 DVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRQPIPKARQP 71

RESULT 4
US-10-467-000-1
; Sequence 1, Application US/10467000
; Publication No. US20040067486A1
; GENERAL INFORMATION:
; APPLICANT: De Francesco, Raffaele
; APPLICANT: Migliaccio, Giovanni
; APPLICANT: Pagnessa, Giacomo
; TITLE OF INVENTION: HEPATITIS C VIRUS REPLICONS AND REPLICON
; FILE REFERENCE: ITR0003P
; CURRENT APPLICATION NUMBER: US/10/467,000
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: PCT/EP02/00526
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/263,479
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Con 1 HCV isolate nucleic acid
US-10-467-000-1

Query Match 26.7%; Score 51; DB 15; Length 3010;
Best Local Similarity 100.0%; Pred. No. 2.5e-37;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRQPIPKARQP 71
Db 21 DVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRQPIPKARQP 71

RESULT 5
US-10-651-165-217
; Sequence 217, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651.165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 217
; LENGTH: 319
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-217

Query Match 26.2%; Score 50; DB 15; Length 319;

```
Best Local Similarity 100.0%; Pred. No. 3.1e-37;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRGRQPIPKAR 70
DB 21 DVKPPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRGRQPIPKAR 70

RESULT 6
US-10-431-587-1
; Sequence 1, Application US/10431587
; Publication No. US20040072267A1
; GENERAL INFORMATION:
; APPLICANT: BIORAD PASTEUR
; TITLE OF INVENTION: Method for simultaneously detecting an antigen of, and an antibody
; FILE REFERENCE: BET 03P0456
; CURRENT APPLICATION NUMBER: US/10/431,587
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: FR 0205808
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-431-587-1

Query Match 25.7%; Score 49; DB 15; Length 75;
Best Local Similarity 100.0%; Pred. No. 7.7e-37;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRGRQPIPKAR 69
DB 21 DVKPPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRGRQPIPKAR 69

RESULT 7
US-09-921-397-3
; Sequence 3, Application US/09921397
; Patent No. US20020151484A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a
; FILE REFERENCE: B4809A - JAZ
; CURRENT APPLICATION NUMBER: US/09/921,397
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: EP 00402225.7
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-921-397-3

Query Match 25.7%; Score 49; DB 9; Length 77;
Best Local Similarity 100.0%; Pred. No. 7.9e-37;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRGRQPIPKAR 69
DB 8 DVKPPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRGRQPIPKAR 56

RESULT 8
US-09-758-308-1
; Sequence 1, Application US/09758308
; Patent No. US20020090607A1
```

```
; GENERAL INFORMATION:
; APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
; TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
; FILE REFERENCE: 14114.0349U2
; CURRENT APPLICATION NUMBER: US/09/758,308
; CURRENT FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/092,339
; PRIOR FILING DATE: 1999-07-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-09-758-308-1

Query Match 25.7%; Score 49; DB 9; Length 91;
Best Local Similarity 100.0%; Pred. No. 9e-37;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRGRQPIPKAR 69
DB 21 DVKPPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRGRQPIPKAR 69

RESULT 9
US-09-891-983A-16
; Sequence 16, Application US/09891983A
; Publication No. US20030108858A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Shah, Dinesh O.
; APPLICANT: Dawson, George A.
; APPLICANT: Muerhoff, A. Scott
; APPLICANT: Jiang, Lily
; APPLICANT: Gutierrez, Robin A.
; APPLICANT: Leary, Thomas P.
; APPLICANT: Desai, Suresh
; APPLICANT: Stewart, James L.
; TITLE OF INVENTION: Methods For The simultaneous Detection
; FILE REFERENCE: 6821.US.O1
; CURRENT APPLICATION NUMBER: US/09/891,983A
; CURRENT FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-09-891-983A-16

Query Match 25.7%; Score 49; DB 10; Length 94;
Best Local Similarity 100.0%; Pred. No. 9.3e-37;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKPPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRGRQPIPKAR 69
DB 15 DVKPPGGQIVGGVYLLPRGRLGVRATRKTSERSQPRGRQPIPKAR 63

RESULT 10
US-10-173-480-16
; Sequence 16, Application US/10173480
; Publication No. US20030152948A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Shah, Dinesh O.
; APPLICANT: Dawson, George A.
; APPLICANT: Muerhoff, A. Scott
; APPLICANT: Jiang, Lily
; APPLICANT: Gutierrez, Robin A.
```

```

; APPLICANT: Leary, Thomas P.
; APPLICANT: Desai, Suresh
; APPLICANT: Stewart, James L.
; TITLE OF INVENTION: METHODS FOR THE SIMULTANEOUS DETECTION
; TITLE OF INVENTION: OF HCV ANTIGENS AND HCV ANTIBODIES
; FILE REFERENCE: 6821.US.P1
; CURRENT APPLICATION NUMBER: US/10/173,480
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/891,983
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Variant - HCV-Core Recombinant
US-10-173-480-16

Query Match          25.7%; Score 49; DB 14; Length 94;
Best Local Similarity 100.0%; Pred.No. 9.3e-37;
Matches 49; Conservative 0; Mismatches 0; Indels 0
QY 21 DVKPPGGQIVGGVYLLPRRGRPLGVRAATKTSERSQPRGRQFIPKAR 69
      |||||
Db 15 DVKPPGGQIVGGVYLLPRRGRPLGVRAATKTSERSQPRGRQFIPKAR 63

RESULT 11
US-10-753-910-16
; Sequence 16, Application US/10753910
; Publication No. US20040185436A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Shan, Dinesh O.
; APPLICANT: Dawson, Greoge A.
; APPLICANT: Muerhoff, A. Scott
; APPLICANT: Jiang, Lily
; APPLICANT: Gutierrez, Robin A.
; APPLICANT: Leary, Thomas P.
; APPLICANT: Desai, Suresh
; APPLICANT: Stewart, James L.
; TITLE OF INVENTION: METHODS FOR THE SIMULTANEOUS DETECTION
; TITLE OF INVENTION: OF HCV ANTIGENS AND HCV ANTIBODIES
; FILE REFERENCE: 6821.US.P1
; CURRENT APPLICATION NUMBER: US/10/753,910
; CURRENT FILING DATE: 2004-01-07
; PRIOR APPLICATION NUMBER: US/10/173,480
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/891,983
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Variant - HCV-Core Recombinant
US-10-753-910-16

```

```

; Sequence 192, Application US/09899046
; Publication No. US20030008274A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/899,046
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 192:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-899-046-192

Query Match 25.7%; Score 49; DB 10; Length 96;
Best Local Similarity 100.0%; Pred. No. 9.5e-37;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFPGGGQIVGGVILLPRRGPELGVRAKTKTSQPRGRROPIPKAR 69
Db 21 DVKFPGGGQIVGGVILLPRRGPELGVRAKTKTSQPRGRROPIPKAR 69

RESULT 13
US-09-878-281-192
; Sequence 192, Application US/09878281
; Publication No. US20030032005A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/878,281
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 192:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-878-281-192

```

RESULT 12
US-09-899-046-192

```
US-09-873-224-192
; Sequence 192, Application US/09873224
; Publication No. US20030064360A1
; GENERAL INFORMATION:
; APPLICANT: <Unknown>
; TITLE OF INVENTION: New sequences of hepatitis C virus
; genotypes for diagnosis, prophylaxis and therapy.
; NUMBER OF SEQUENCES: 270
; CORRESPONDENCE ADDRESS:
; STREET: Industriepark Zwijnaarde 7, box 4
; CITY: Ghent
; COUNTRY: Belgium
; ZIP: B-9052
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/873,224
; FILING DATE: 05-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Innogenetics sa.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 00 32 9 241 07 11
; TELEFAX: 00 32 9 241 07 99
; INFORMATION FOR SEQ ID NO: 192:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 192:
US-09-873-224-192

Query Match 25.7%; Score 49; DB 10; Length 96;
Best Local Similarity 100.0%; Pred. No. 9.5e-37;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFFGGGQIVGGVLLPRRGRLGVRATRKTSERSQPRGRQPIPKAR 69
Db 21 DVKFFGGGQIVGGVLLPRRGRLGVRATRKTSERSQPRGRQPIPKAR 69

RESULT 15
US-09-756-875-8
; Sequence 8, Application US/09756875
; Patent No. US20020150990A1
; GENERAL INFORMATION:
; APPLICANT: PIKE, IAN
; TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESS: Suite 701-E Columbia Square
; STREET: 555 13th Street, N. W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U. S.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/756,875
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 08/259,721
; FILING DATE: 29-AUG-1994
; APPLICATION NUMBER: PCT/GB93/00410
; FILING DATE: 26-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: ERNST, BARBARA G.
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1808-157A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)783-6040
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-756-875-8

Query Match 25.7%; Score 49; DB 9; Length 97;
Best Local Similarity 100.0%; Pred. No. 9.5e-37;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKFFGGGQIVGGVLLPRRGRLGVRATRKTSERSQPRGRQPIPKAR 69
Db 21 DVKFFGGGQIVGGVLLPRRGRLGVRATRKTSERSQPRGRQPIPKAR 69

Search completed: October 30, 2004, 02:03:53
Job time : 50 secs
```

